DOCUMENT RESUME

ED 180 768 SE 029 395

AUTHOR	McRobbie, Campbell J., Ed.; Power, Colin N., Ed.
TITLE	Research in Science Education, Volume 8. Proceedings of the Annual Conference of the Australian Science Education Research Association (9th, Mount Gravatt College of Advanced Education, Brisbane, Queensland, May, 1978).
INSTITUTION	Australian Science Education Research Association.
PUB DATE	78
NOTE	211p.; For Volume 7, see ED 168 815
AVAILABLE FROM	Australian Science Education Research Association, School of Education, Flinders University, Bedford Park, 5042, Australia ($A 10.00)
EDRS PRICE	MF01/PC09 Plus Postage.
DESCRIPTORS	Curriculum Evaluation; *Educational Research; Elementary School Science; Elementary Secondary Education; Higher Education; *Inquiry Training; *Questioning Techniques; *Science Education; Scientific Enterprise

ABSTRACT

Presented are the 21 papers which constitute the proceedings of the ninth annual conference of the Australian Science Education Research Association at a meeting in May, 1978. Papers range from philosophical analysis of the language of science teaching to experimental research into the effects of different laboratory skill development strategies, from cognitive development in primary school children to analyses of teaching in university microbiology classes, and from evaluations of facilities and curricula to the rough theatre of the science classroom. Most of the studies fall within the psychometric paradigm and have used its methodologies, although one or two have used philosophical methods of analysis. (Author/PB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED180774

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Colin N. Power

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

SE 029 395

RESEARCH
IN
SCIENCE EDUCATION

VOLUME 8

Proceedings of the Ninth Annual Conference of the Australian Science Education Research Association, Mount Gravatt College of Advanced Education, Brisbane, Queensland, May, 1978.

Edited by: *Campbell J. McRobbie, Mount Gravatt College of Advanced Education*

General Editor: *Colin N. Power, Flinders University*

Editorial Board: Arthur Lucas, Flinders University
Max Maddock, University of Newcastle
Leo West, Monash University
Richard White, Monash University

Business Manager: Jim Butler, University of Queensland

AUSTRALIAN SCIENCE EDUCATION RESEARCH ASSOCIATION

All general correspondence concerning this publication should be addressed to Professor Colin N. Power, School of Education, Flinders University, Bedford Park, 5042, Australia.

Library subscriptions to the Business Manager, Dr J. Butler, Education Department, University of Queensland, St. Lucia, 4067, Australia.

Copyright © 1978 by Australian Science Education Research Association.
Published by Australian Science Education Research Association.
Printed by the University of Queensland Printery.

CONTENTS

Page

Preface (v)

Barry C. Newman
The scientific status of any so called science 1

J.E. Butler
Epistemology in the language of the science classroom 11

Rod Fawns
The rough theatre of science teaching in the compulsory years rationality & artistry 25

G. Wayan Seregeg
Some implications of the introduction of enquiry methods to the integrated science curriculum in Indonesia 35

Elizabeth H. Hegarty
Levels of scientific enquiry in university science laboratory classes: Implications for curriculum deliberations 45

D.A. Mitchell
The production of criteria for evaluating science curriculum materials with specific reference to criteria for inquiry and the application of these criteria to chemistry curriculum materials 59

John Dekkers
The effects of junior inquiry science programs on student cognitive and activity preferences in science 71

A. John Rentoul and *Barry J. Fraser*
Measuring perceptions of inquiry and open learning environments 79

Adrian Fordham
The interaction of student characteristics and science teaching on student perception of the learning environment 89

John Ainely
Science facilities and variety in science teaching 99

John M. Owen
The role of linkage mechanisms in promoting educational change 110

James M. Richmond
Some outcomes of an environmental knowledge and attitudes survey in England 119

M.N. Maddock
An exploratory study in school assessment: The relationship between achievement and attitude scales 127

Contents - contd.

Page

W.F. Beasley
Laboratory psychomotor skill development using physical and mental practice strategies 135

Lawrence G. Little and *Rob. W. de M-Maclay*
Skills and skills tests in experimental physics 145

Miles A. Nelson
Questioning questioning Part 1 157

David J. Symington
Primary school pupil's ability to see investigable scientific problems in everyday phenomena: The teacher's role 167

P.Lynch, P. Benjamin, T. Chapman, R. Holmes, R. McCammon, A. Smith, R.Symons
Scientific language and the high school pupil: Part 1 175

R.D. Linke and *M.I. Venz*
Misconceptions in physical science among non-science background students 183

P.P. Sedgwick, R.D. Linke and *A.M. Lucas*
A comparison of changes in children's concepts of life with the development of relevant criteria in Australian science curriculum materials 195

Graham W. Dettrick
The performance of 9 to 12 year old children on primary, secondary and multiplicative classification tasks: Implications for teaching and curriculum design 205

Contributors 213

6

PREFACE

This volume contains the papers presented at the 9th Annual Conference of the Australian Science Education Research Association held at Mount Gravatt College of Advanced Education in May 1978. The papers reflect the diversity of interests of A.S.E.R.A. members; ranging from philosophical analysis of the language of science teaching to experimental research into the effects of different laboratory skill development strategies; from cognitive development in primary school children to analyses of teaching in University microbiology classes; from evaluations of facilities and curricula to the rough theatre of the science classroom.

Most of the studies reported fall within the psychometric paradigm and have utilised its methodologies. One or two have utilised other (e.g. philosophical) methods of analysis; and in a few papers, there are indications of a willingness to consider methodological alternatives, to tackle hitherto untouched problems and to adopt fresh perspectives. As a consequence, at least three or four of the papers presented represent a challenge not only to the research traditions of science education, but also to science teachers, curriculum developers, and teacher educators.

The continuing interest of members in problems relating to the nature of science and science education is evident in the first two papers. Science education research workers might well like to apply Barry Newman's classification to their efforts. Might the research be described as less than science, or even worse as unknown and unimportant science? If so, what might be done? Jim Butler's paper could not be described as unimportant, but it does raise the question of whether school science is a limited or unsuccessful science if not a misleading one. Rod Fawns paper is undoubtedly unique but it also suggests that we, as teachers and researchers, may need to be more adventurous in our attempts to teach and to study science teaching. Several papers deal with the gap between the objectives of promoting inquiry, psychomotor skills, and affective development, and the reality. Collectively, these papers yield fresh insights and outline some promising tools which might be useful in studying the discrepancies identified. It is encouraging to see that more often than not, a careful study is being made of the patterns of behaviour, the nature of the learning environment, or the curriculum materials used.

Throughout the Conference papers the Australian bias persists, albiet spiced with the delightful flavour of South East Asian analyses of inquiry, a pragmatic look at environmental knowledge and attitudes in the United Kingdom and at psychomotor skills in the United States of America. Once again we find that certain perspectives are missing. For instance in the A.S.E.R.A. Conferences held to date, no paper has undertaken an historical analyses of any aspect of science teaching in Australia. Relatively few science teachers attend A.S.E.R.A. Conferences and rarely, if ever, have science teachers dared to report on their attempts to inquire into problems faced by them. At several A.S.E.R.A. Conferences, workshops (not reported here) have looked at the problems of dissemination and suggestions have been made as to ways in which the apparent gulf between science education research and practice might be narrowed. But, while we have studies of the diffusion of curriculum innovations (cf. John Owen's paper), research on the dissemination of research is all too rare.

Colin Power
General Editor

Cam McRobbie
Editor

Research in Science Education, 1978, 8, 1-9

THE SCIENTIFIC STATUS OF ANY SO-CALLED SCIENCE

Barry C. Newman

Introduction

Recently, in a British Broadcasting Corporation interview (1978), Sir George Porter, Director of the Royal Institution, upon being asked when it was that he first became interested in science, replied that it was when he was about one year old. In the same interview he was also asked what his advice would be to those who were thinking of taking up a career in science. In his reply to this question he said he would advise them not to go into any of the social sciences as these really were not sciences. What his remarks imply, among other things, is that things called science are held in esteem and that consequently it is of importance to deny or affirm that any so-called science is indeed science.

Simply "yes" or "no" answers to the question "Is this science?" are, however, sometimes too easily, and perhaps too readily, given. On other occasions no answer is readily forthcoming at all because either of the alternatives "yes" or "no" seems so obviously deficient. The main purpose of this paper is to provide a scheme whereby a judgment on such a question can be given, when "yes" or "no" seems inappropriate, that is more informative than such simple answers. Furthermore, it is intended that the scheme will allow some of the esteem associated with science to be attached to something which might not be held to be 'good, complete, unqualified science', i.e. 'full-blooded science'.

To develop this scheme it will be necessary to examine first 'Science' and then 'Less than Science'. An analysis of 'The Study of Science' should then help in assessing the nature of any particular study of science and the status of the science so studied. Finally, some attention will be given to 'The Judgment' one makes in any specific situation.

Science

Obviously, what one understands by the nature of science is a fundamental area of concern for the purpose of this paper. The whole case for identifying whether Primary Science, Secondary Science or any other so-called science is science or not or in what way it is something less than science, rests upon the position one adopts on this issue.

If, for instance, it is claimed that the nature of science, for one reason or another, cannot be determined to any extent, then we are left with an insurmountable problem. However, we will still have to justify choosing to use the word 'science' in one context but not in another, if we choose to use the word at all. In this paper it will be assumed that we still want to use the word 'science' and with some discrimination as to its use. A stance, therefore, will be adopted on the nature of science, both for the purpose of developing this paper and because presumably we do believe we can say something definitive about the nature of science.

Now the author is not unaware of some of the problems associated with trying to describe the scientific enterprise and elsewhere (Newman, 1977) indicates something of the range of viewpoints that is available. None the less a particular viewpoint in this paper will be expressed, a viewpoint, however, that is capable of some defence. It ought to be borne in mind, however, that if a different position on the nature of science is adopted it does not necessarily follow, as is indicated below, that the whole scheme for handling the question "Is this science?"

breaks down.

In this paper, science-'full-blooded science' is regarded as being characterised by four *functions* and one *aim* with the *aim* being *realised*. The four functions are:

The Collection of Data
The Construction of Theories
The Testing of Theories
The Replacement of Theories

and the aim is that of-

Increasing the Explanatory Power of Theories

The main defence for this position, though brief, is made by appealing to the philosophers of science, Popper, Kuhn, Lakatos, Feyerabend and Althusser (for a brief introduction to the views of these philosophers, see Newman 1975 and 1977).

The Functions of Science

Each of these philosophers sees a place in his view of science for each of the functions mentioned above, though the actual model that each has is substantially different in each case.

Althusser (1969) would undoubtedly see the collection of data as work done on some raw material which, together with other raw material, is transformed, during scientific practice, into some of the products of such practice. For Popper (1972*a*) such practices external to scientific practice are not important. He would see data collection as having its main significance in the attempted falsification of theories. Kuhn (1970) would view the collection of data as part of a normal science puzzle solving process or as part of an attempt to support a paradigm under attack. Lakatos (1970) would see the collection of data, together with alternative theories as means for replacing theories by more embracive ones. Feyerabend (1975), whose main concern is the proliferation of theories without methodological guidance or restraints, would allow data collection likewise to proceed without methodological direction.

Each has a place for the collection of data though each would see the part played by it differently. Likewise, it could be established that each sees a role for the other functions mentioned, though again, the role taken is seen to be somewhat different in each case. No attempt will be made to establish their interest in these other functions, for it is even more obvious than their concern for the collection of data that they have a place for theory construction, theory test and theory replacement.

Certainly, each has emphases which would not be satisfied by the statement above on 'full-blooded science', on its own. Feyerabend and Kuhn see an important part in the progress of science for incommensurable theories. Kuhn also places stress on scientific communities. Althusser has an important role for the labour which is involved in the production of scientific theories together with the relationship of the scientific enterprise with other practices of a society. Lakatos' main emphasis is on research programmes of theory development with each theory having a hard core and a protective belt. Popper's emphasis is on falsification with corroboration being a consequence when falsification fails. For the purpose of this paper it is sufficient, however, that these philosophers see a place and an important place generally for the four functions mentioned.

The Aim of Science

The aim of science- stated as, increasing the explanatory power of theories, is introduced, in part, because a notion of explanation seems implicit, if not explicit, in many analyses of the nature of science. Popper (1972*b*) explicitly refers to it claiming that the aim of science is seeking explanations. Kuhn (1970) seems to be referring to explanations in his notion of the existence of puzzles and their possible solution. Lakatos' progress in a research programme (1970) can be interpreted as progress enabling 'better' explanations to be given of more phenomena. Explanation seems implicit in Feyerabend's ideas of 'progress' and 'better' theories (1975).

In addition, however, the idea of explanation is introduced in order to be able to differentiate between scientific activities and say, the activities of a group of thieves who being determined to steal more effectively, collect data, build theories, etc. (see Feyerabend, 1970). It is having the aim of increasing explanatory power which serves to distinguish between such activities.

That this aim should be realised is also included as characteristic of science, as it were, at its best, because it seems so often implied in notions of progress in the scientific endeavour. If the aim or aims of science, whatever they were considered to be, were never or only rarely realised, most of the esteem presently associated with science would probably not exist.

Alternative Models for Science

It is quite likely that the model proposed as characteristic of science is unsatisfactory for some, for whatever reasons. Some might argue that the notions of 'the scientific community', 'inductivism', 'falsification of theories' or any one or more of a host of ideas should be given a prominent and appropriate place in, say, the functions of scientific endeavour. Furthermore, it might be maintained that the aim of science has to do with prediction or control of our world, together with or instead of explanation of it. Some might suggest that as well as *function* and aim, science has to be characterised by *Area*, that is, that which is to be investigated, a reference to the type of data and the substance of the theories. An appeal to 'area' might be made in order to distinguish science from, say, history. Incidentally, it is to be noted that introducing the idea of 'area' could be of help in answering a question along the lines of "Is this biology?" or similar. The additional notion of 'area' is introduced to delineate biology from other sciences, and the scheme that follows with some modification and perhaps further elaboration, is still applied to enable one to give an informative answer to that question.

The point to be made is that it might not matter very much what one's model of science is, for the scheme to be outlined to be applied. Before the scheme can be developed, however, one has to have some idea of what one is prepared to call 'good, complete, unqualified science'. A model of such science, somewhat different to the one given here, might mean some alterations, even drastic alterations to the scheme which follows. However, the nature of those alterations, hopefully, in many cases, would be apparent once the scheme based on the model adopted here is understood.

Less than Science

Perhaps on more occasions than not, an activity may be carried out, which though seemingly scientific may not be regarded worthy of treating as 'good, complete, unqualified science'. Such science, in this paper, will be referred to as 'science less than science'. Seven types of 'science less than science' are outlined and brief mention is made of one type of 'non-science less than science' referred to as 'pre-science'.

1. Limited Science

This categorisation turns on the notion that in order for an activity to be called science at all, at least one of the functions referred to in the model of science above needs to be evident and that, in addition, this function needs to be carried out for the purpose of increasing explanatory power. For example, where a theory is constructed with the intention of increasing the explanatory power that one has, but where no testing of the theory takes place, no data is collected or where no theory is replaced (there being perhaps only a comparison of theories, or no theory to begin with) and whether the aim is realised or not, then such an activity is referred to as an example of limited science. True, it may be regarded as very limited science, but science none the less. The extent to which, and the way in which it is limited can be pointed out in describing it as limited science.

2. Unsuccessful Science

Whether limited or not, a scientific activity can be further classified as successful or unsuccessful depending on whether or not the aim of seeking greater explanatory power is realised. If it is not realised, then on that count alone, it is 'science less than science' - the science has been unsuccessful. One can fail to achieve the aim if no theory is forthcoming or if the theory produced simply does not have excess explanatory power over a theory one already has. A scientific activity then, can be characterised as unsuccessful in at least one of two senses. Furthermore, it should be noted, that in order for any scientific activity to be successful the activity cannot be too limited. Some type of theory construction and theory replacement must at least be carried out.

3. Insignificant Science

A scientific activity can be considered to be relatively insignificant if one or more of those aspects which characterise it as scientific is of poor quality. For example, the data collection could be careless, the tests applied could lack power, the theory or theories constructed could embrace few phenomena, or the more powerful explanation achieved could be only minimally more powerful. Obviously, a scientific activity can be more or less significant or insignificant along various dimensions.

4. Immature Science

In a Kuhnian sense (1970), a scientific activity can be considered to be immature or mature depending upon the extent to which the members of that group having the same area of interest hold to opposing theories or sets of theories, in seeking to explain the same phenomenon. It is mature science if a significant number of the significant theories appealed to are held in common by a significant number of the group working in the same area. A scientific activity may not be clearly either immature or mature, but might be described as more or less mature or immature, with respect to that fraction of the group having the same area of interest, making appeal to the same theories, that fraction of the theories applied to the same area being held by the group, and the degree of significance, with respect to the area being investigated, of the theories so held. It should be noted that the notions of mature or immature science diminish in their applicability as the size of the group working in the same area diminishes. If only one person is working in an area, the notions do not apply at all.

5. Non-unique Science

Now references to mature and immature science above, only apply to the activities of one group whose members are acquainted to some extent with each other's activities. Sometimes, however, the one area may be being investigated, with much the same activities being carried out by groups (or individuals) that are not aware of each others work. In this case they are said, by someone who becomes acquainted with both groups and their activities, individually to be doing non-unique science. Regardless of the reason, be it political, economic, religious or whatever, this lack of acquaintance of one group with another group's similar activities is not a good reason for discrediting the worth of such activities. They have simply been carried out more than once. In fact, when viewed as one activity, their joint efforts may be seen to be of greater worth than either effort alone, because of the greater confidence that often attends replication.

Even when one group or individual is preceded by another in accomplishing the same scientific feat, the honour commonly ascribed to the one who first carries out his work or perhaps first publishes it, is an honour seemingly given to he who comes first in some sort of race. The honour is not given because the science is superior. The later scientific activity is simply an innocent duplication, unless it is claimed that the activity carried out first was in some way more difficult. In this case, for the activity carried out later, the category 'non-unique science' seems to some extent inappropriate, and perhaps the category 'insignificant science', along one or more dimensions, is more appropriate.

Strictly speaking, non-unique science is not 'science less than science' except in so far as the non-unique science carried out second or whatever, is given a later place in the race. For this reason, where a time difference - normally the case, is involved, only the second and any subsequent similar activities should be labelled instances of 'science less than science'.

6. Unknown Science

When the results of a scientific activity are only poorly communicated to others working in the same area or even not communicated at all, then such scientific activity is referred to as 'unknown' or relatively 'unknown science'. It may be only a little limited, quite successful, rather significant and unique, but if not communicated to others or only poorly done so, it is generally regarded as suffering somewhat as a consequence. As with non-unique science it is strictly speaking, by virtue of its unknown or relatively unknown quality alone, not 'science less than science', except in this case, in so far as it would have been regarded more highly if it had been well publicised.

7. Unimportant Science

Any scientific activity can be judged to be more or less important or unimportant by appeal to a number of criteria. Its importance can be judged, for example, with respect to the scope of its area of concern, the nature and extent of the further scientific research it generates or the extent to which it aids in meeting significant human needs. As with non-unique science and unknown science, unimportant science, with respect to the quality of importance alone, is also not strictly 'science less than science' except in this case, in so far as any scientific activity is considered to have greater worth, the greater its importance.

Pre-science

Where some activity is carried out that has or could have some link with some genuine science activity, but is in itself not a scientific activity, such activity is described as pre-scientific. Some examples of pre-science are: the collection of data without any thought of seeking to explain some phenomenon but carried out as part of some scientific training, the construction of equipment that is intended to be used for some scientific endeavour, the acquisition of the skills required to handle such scientific equipment in order to enable one to engage in some scientific work. It is necessary that there be at least some link of intention between what is done and a genuine scientific activity before the word 'science' can be used, in describing what is done, as pre-science. It is not science itself, but because of its link with science, the activity is given something of the esteem of science.

The Study of Science

It could well be the case that much of what is termed 'doing science' in primary, secondary and tertiary science is at best, the study of science and it is not necessarily the case that the study of science implies being engaged in genuine scientific activity. It is important, therefore, to examine what can legitimately be called such a study and to determine something of the relationship that might exist between the study of science and the carrying out of genuine scientific work. In what follows, the science component of the study of science and the study component of the study of science are examined somewhat independently.

The Science Component of the Study of Science

Because of the position already adopted on the nature of science, there are certain things that the science studied cannot be, in order for such study to be properly called the study of science. It cannot be, properly speaking, the history of science or the philosophy of science, though these may be used to assist in the study of science, and though these may be studied, somewhat unintentionally, as a consequence of the study of science. It cannot be what scientists do, unless what one means by this is what scientists do as scientists. In this case, it can only be what science is, that is science, as previously stipulated.

Just as science can be more or less limited, successful or unsuccessful, more or less significant, more or less mature or immature, unique or non-unique, more or less important and more or less known, so can the science studied be of that nature. However, usually there must be something of an 'out of date' character to the science studied simply because it is studied. It has already been done. Admittedly, it could be possible for a student or group of students to view it while it was being done.

Because of the nature of educational enterprises, any science studied would generally be of a very limited nature, for reasons of economy of time. Also, it is probably the case that most science studied would be rather successful, significant, mature, unique, well known and according to certain criteria, quite important. If the study of science is intended to make the student aware of the actual scientific enterprise, as practised, then there should be more of the unsuccessful, insignificant, immature, non-unique, poorly known and even unimportant science introduced.

The Study Component of the Study of Science

Studying science in no way implies actually engaging in scientific activity. Studying science does not prevent the carrying out of scientific work but it is not necessarily carried out as a result. One could study the genuine doing of science by just reading about it or by having someone tell one about it, without having to go through any of the motions.

On the other hand, a teacher or lecturer could decide to have his students studying science by purposefully having them, in varying degrees, actually involved in scientific activities. At one extreme of such activities the students could be copying a procedure of, for example, data collection or theory construction, where there would be little likelihood of their own self-initiated, self-guided activity occurring. It would not be unique science, for it would not be original. However, it would not be non-unique science either, since the students would be working with the knowledge of someone else's activities, which activities would be similar to their own. It has already been stipulated that for a scientific activity to be non-unique, different groups or individuals have to be engaged in similar activities, without the knowledge of each other's work. Would it be science at all? In an important sense they are not doing science. Someone or some number of people, in one way or another had a hand in doing such science, in the past, but a person actually copying what was done is doing simply that, copying the doing of science rather than doing it. Certainly he may be studying science.

At the other extreme of such activities, the general area of investigation might be selected by the teacher or lecturer, or even by the student himself, with the student being given relatively free reign and little guidance as he proceeds to engage in data collection or theory construction, etc. Now this time it might be very likely that the science-like activity would be non-unique. In addition, it would probably be, but not necessarily so, relatively insignificant and unknown and, according to certain criteria, relatively unimportant as well. Could it still be science though? It would certainly not be copying science. It might or might not be limited, though it would probably be more limited than any original science it was duplicating, because of the limited amount of time that could be spent on it. It could be successful or unsuccessful. Since in many cases the students would almost certainly not be aware of any similar activity carried out by others outside of their own educational setting, where a group of them were interested in a similar area and where a significant number of them were aware of each other's contributions, it could eventuate that their activity by itself was mature or immature. The important point is that provided a student or students were not copying science, provided they were attempting to seek a more powerful explanation than any they had (often none to begin with?) and provided they were engaged in at least one of the functions of data collection, etc., then their activity should be described as a scientific one. Even in the case where a general area of investigation is chosen for a student, this is not a reason for its being ruled out as science. In one way or another, areas of research in which members of the officially recognised scientific community work may well often be chosen, to some extent, by people other than those actually engaged in the research. Here then is a situation where though a teacher's or lecturer's prime purpose could be for his students to learn about science, his students could end up engaging in genuine scientific work and, as a consequence, be studying science. Of course, the teacher's or lecturer's main purpose could be for his students to be actually engaging in scientific activity rather than studying it.

The Judgment

Enough has been said now, hopefully, for a judgment to be made on any so-called science as being the study of science, the genuine doing of science, the copying of science or the engaging in some pre-scientific activity or some mixture of these and then to determine whether any science component is more or less limited, successful or unsuccessful, more or less significant, more or less mature or immature, non-unique or unique, more or less known and more or less important.

It is necessary, however, for purposes of clarity, to be specific about the so-called science referred to. One could be making a judgment on an activity of 30 minutes duration or a life-time's work, the results or anticipated results of a term's programme at a tertiary institution, or the results or anticipated results of following a primary science text, or the activities of an individual or a group of people. Any judgments made are judgments concerning that particular point of focus.

Furthermore, in adopting the scheme or something like it, one can choose to make an assessment in very general terms, for example, saying about some activity that it is reasonably significant but very unimportant, or in very specific terms, for example, outlining various criteria used for significance and importance and specifically pointing out how the activity fared according to these criteria.

Finally, as with all judgments, a judgment in this area cannot be but subject to review. This is so not only because new evidence may come to light that causes us to change our judgment but also because the model of science we are prepared to work with and what we understand, for example, by limited science, insignificant science and important science, etc. may well change (nay should?) anyway.

Conclusion

To return to the introduction, was Sir George Porter engaging in genuine scientific activity when one year old? Who knows, but if at that age he was not copying someone else, was seeking greater explanatory power and carrying out at least one of the functions of collecting data, etc., then according to the scheme outlined in this paper, he undoubtedly was. True, it would almost certainly be very limited, very insignificant, very unimportant (but perhaps not to him), non-unique and relatively unknown, but with a fair chance perhaps of being successful. The question of maturity does not apply.

Are the social sciences sciences? One hopes that Sir George, upon reading this analysis would be forced to say yes, although with some glee he might maintain that the only difference between his scientific activity at age one and that of the social sciences, apart from the different area of interest, is that the latter is immature.

References

ALTHUSSER, L. *For Marx,* translated by B. Brewster. Harmondsworth: Penguin, 1969.

BRITISH BROADCASTING CORPORATION. *Profile . . . . Sir George Porter.* BBC Topical Tapes, Tape No.: 8R46P813P, Recorded 17th March, 1978. Broadcasted by the Australian Broadcasting Commission, Sydney, 7.15 p.m., 14 April, 1978.

FEYERABEND, P.K. Consolations for the specialist. In I. Lakatos and A. Musgrave (eds), *Criticism and the growth of knowledge.* Cambridge: Cambridge University Press, 1970, 197-230.

FEYERABEND, P.K. *Against method.* London: NLB, 1975.

KUHN, T.S. *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago Press, 1970.

LAKATOS, I. Falsification and the methodology of scientific research programmes. In I. Lakatos and A. Musgrave (eds), *Criticism and the growth of knowledge*. Cambridge: Cambridge University Press, 1970, 91-196.

NEWMAN, B.C. Scientific methodology: a Feyerabendian view. *Research in Science Education,* 1975, 5, 99-111.

NEWMAN, B.C. The nature of science problem. *Research in Science Education.* 1977, 7, 1-7.

POPPER, K.R. *The logic of scientific discovery.* 6th impression (revised). London, Hutchinson, 1972a.

POPPER, K.R. *Objective knowledge.* Oxford: Oxford University Press, 1972b.

Research in Science Education, 1978, 8, 11-23

EPISTEMOLOGY IN THE LANGUAGE OF THE SCIENCE CLASSROOM

J. E. Butler

> If speech systems are generated or controlled by forms of social relations, then discoveries, however fascinating and well described, cannot be explained without a theory that embraces social relations.
>
> B. Bernstein, 1972

A central focus for recent critics of the schools has been the 'absolutist' and 'objectivistic' conceptions of knowledge which dominate the school curriculum. A related target is the traditional 'transmission of knowledge' model of teaching and learning. 'Reified', 'out-there', 'taken-for-granted' are but a few of the labels attached to the manner in which knowledge is treated in such a positivistic epistemology. The analyses of these critics have centred on the tendency of teachers to treat knowledge as somehow absolute and independent of themselves and their students, and as if it corresponds directly to Nature and eternal reality. Thus the critics point to the way in which this epistemology minimizes human agency in the production of knowledge. Alternatively, they would argue that teachers should stress man's being active in giving meaning to his world and that knowledge is a human production. The students would then be able to recognize that there is no one 'reality' and thus they will develop a permanently critical approach to knowledge.

The studies reported in this article sought to investigate and explain the view of knowledge which is portrayed in the science classroom. The data base was the language used in science classrooms. Thus the specific purposes of the studies were: (a) to examine the language of the science classroom to see what view of science was being conveyed; (b) to explain how the epistemological necessities involved are produced. However, before this preliminary report is given several germane topics will be reviewed briefly.

Language in the classroom

There is a strong tradition of classroom research based on the spoken language of the teacher and students, (Bellack, 1966; Flanders, 1970; Lundgren, 1972; Barnes, 1976; Young, 1977). Such studies are usually based upon typed manuscripts of audio-recordings made during the lesson. The study reported here is partly based on such transcripts, but examines them from a theoretical perspective that differs from those used by any of the above except Young (1977).

I accept as a starting point the twofold function of classroom language as identified by Barnes (1976). The first function of classroom language is as a medium of communication. In the analysis of communication it is usual to distinguish between the semantics proper or reference, and pragmatics. Reference indicates the objective characteristics of the communication whereas pragmatics deals with the effect of the symbols on the speaker and hearer. It will be the pragmatics of the classroom communication that will be examined here. Specifically, the main interest is in the effect that the classroom language has on the manner in which the students experience scientific knowledge. It is assumed that there is a relationship between the students' view of knowledge and the form of communication in the classroom language.

The second function of classroom language is that it is a means of learning. Barnes (1976) argues convincingly the notion that the changing and formulating of knowledge by the act of speaking is an essential function of the language in the classroom. This language function is not central to this paper.

Barnes (1976) deals exclusively with the spoken language during teacher-student interaction. For my purposes I will include in the concept of classroom language the written language of the science textbook. The analyses of the spoken and written languages will be similar in many ways but differences will be noted below.

The methods employed in this paper follow the tradition of Barnes (1976). However my purpose is different. The only related work was published by Munby (1976), but he used a different theoretical viewpoint and dealt only with the language of textbooks.

The origins of the criticism of school science

Two newly developed perspectives (Kuhn, 1970; Young, 1971) have focussed criticism on the way knowledge is portrayed in the science classroom. These perspectives have thrust school science to centre stage in the curriculum debate, and in their twin lights a positivistic imposter has been exposed.

A revolution in the philosophy of science has been brought about by, *inter alia*, Kuhn (1970), Feyerabend (1970) and Toulmin (1972). The traditional positivistic philosophy has been characterized by an ahistorical approach with a heavy emphasis on logic, on the principle of empiricism and on formal analyses. As such it is entirely separated from any activity of the scientist *qua* human being. When this view is pressed we find a tacit ideology: the science of science should be science. But this leads to the exclusion of all that is importantly human from the account, since most of the human activity slips the net composed of the warp of basic descriptive statements and woof of logical machinery.

Over against this abstractive approach the new contender proposes placing the scientist at the centre of the social-cognitive nexus known as science. It proposes studying historically and sociologically the lives of scientists, and from such a rational account to turn to the task of asking in what measure an abstract notion of science can be rescued from the community activities of scientists. The detailed systematic development of this approach lies in the future; what Kuhn has done has been to reveal the great power of this reorientation.

A revolution has also been achieved by the advent of the 'new sociology of education' (Young, 1971). Representatives of the recent orientation in the sociology of education attack the 'reification' of the content of the curriculum and describe the experience of school science as 'alienation' (Young, 1977). They concentrate their studies upon the social processes whereby particular conceptions of school knowledge are legitimated and sustained. According to Gorbutt (1972) this revolution draws its intellectual strength from Marx, C.W. Mills, Schutz, Berger, Garfinkel and others. The basic agreement of these writers is in their critique of a prevailing 'positivist' ideology. This is the legacy they have bequeathed to the 'new' sociology.

Language and thought

It is appropriate at this stage to draw together the ideas canvassed so far. This can be done by comparing the positivists and their critics when they discuss the extent to which forms of language leave their mark on the way in which knowledge is experienced.

The positivists claim (Hooker, 1975) that human beings acquire language in two quite distinct ways. They derive the empirically significant component directly from specific sensory experiences, this is the first mode. The remainder, which is that part of language devoid of empirical content, consists of the logical and social-emotive components which are acquired through the non-rational process of socialization in the human community. This latter mode produces a language that consists of only pseudo terms and pseudo sentences which have no cognitive content. On the other hand, for the positivists any simple sentence stated as an immediate serious report of sensory experience is certainly *true*, therefore the language derived by the first mode – the language of science – conveys cognitively meaningful true knowledge. In summary, according to the positivists scientific language is the bland vehicle for the expression of the contents of thought and perception, neither of which is influenced by the language of science.

The critics of the positivists deny that people derive any of their language by the first mode as described above. According to this new view (Mills, 1963) a theory of mind is needed which conceives social factors as intrinsic to the whole of our mentality. It is only by using his community's language that a person can think and communicate. Language is socially built and maintained, and when the categories of a language are acquired, the ways of the group are acquired. All behaviour and perception, all logic and thought come within the control of a system of language. "Vocabularies socially canalize thought." In summary, according to Mills thinking influences language very little, but thought and perception are largely determined by a socially derived language.

Results of analysis of classroom language

Two methodologies were used, in a limited way at this stage, to investigate the view of science embodied in the language of the science classroom.

1. Linguistic Analysis

This methodology is derived from Gusfield (1976) who analyzed linguistically one of a set of published scientific papers, to support his hypothesis that the paper portrayed its conclusions in a positivistic mode. Gusfield identified a range of writing techniques by which the author of the article gave his results the appearance of absolute truth characteristic of the positivist position. The results presented below are examples of these same techniques found in a popular senior secondary Physics textbook (Mayfield *et al.*, 1973), and in the transcripts of Grade 8 science lessons in some Brisbane schools.

According to the positivistic view scientific language is a non-distorting medium by which the external world is reported and communicated. Gusfield calls this the "windowpane" theory of language. It insists on the intrinsic irrelevance of language when scientific knowledge is transmitted, for example, in the science classroom. The result of communicating concepts and data is to enable the recipient to see the natural world as it really is. Therefore, the language and form of the communication must be chosen so that it closely approximates a pane of clear glass, as Gusfield puts it "The style of non-style is itself the style of science". Gusfield identifies the key aspects of this non-style as the stragegy which places the scientific method as the central actor and then insists on the externality of data. These aspects are usually manifested by the author not letting any part of himself intrude into the language, and by writing in a way which emphasizes the inevitability of the scientific method and thus the passive role of the scientist. The following excerpts illustrate these same strategies in the language analysed in this study:

(a) Externality of data

"The results shown in Fig. 3.6 are typical of this experiment and show that the acceleration A is directly proportional to the unbalanced force F."

(Mayfield *et al.*, p. 38)

"Similar calculations based on photographs of other collisions show that the action and reaction forces are always equal in magnitude and opposite in direction."

(Mayfield *et al.*, p. 44)

Right. We're rubbing it. I want you to watch what happens. See what happened to the paper? What does that remind you of? What happened then?

P: Magnetism.
T: Reminds you of something that we've done recently. What, Len?
P: Magnets.
T: Yes. You think it's the same thing? Same sort of thing: What's making this attract the paper? Ian?
P: Heat.
T: Hmm?
P: Heat.
T: No, not heat. Now I'm rubbing this. What does that cause? Rubbing causes?
P: Friction.
T: Hands up.
P: Friction.
T: Friction and what does the friction cause in this particular case? Yes?
P: Heat.
T: No. It can cause heat and it can cause wear but it's also causing something else in this case. Yes?
P: Electricity.
T: Yes. It's charge. It's causing this rod to become charged and now it's able to pick up fairly big pieces of paper.

(Grade 8 lesson)

(b) Passivity of the Scientist using the Scientific Method

"(Newton's) study of similar collisions and his observations of the action and reaction forces involved led him to make an important generalization

Newton's Third Law."

(Mayfield *et al.*, p. 44)

"We are lead to the following conclusion:

A constant unbalanced force produces a constant acceleration."

(Mayfield *et al.*, p. 37)

P: Like charges repel, unlike charges attract.
T: Yes. Tell me what would happen if I brought two negative rods together. Quick. Yes?
P: They repel.
T: Right. Two glass rods together? Yes?
P: Repel.

T: Okay. A glass one and an ebonite one? Yes?
P: Attract.
T: An ebonite one and a glass one?
P: Attract.
T: Yes. That proves our law, doesn't it? Do you think it proves it entirely? Okay. Now, I want people to come out here and prove that law for me. I'll have (pause) I want Christine, you had better come out and prove to me that like charges attract. Okay, are you going to prove that like charges attract? Do you know how to prove that? Who could prove that like charges attract. Okay. Off you go. (pause – class laughs). What do you think Jan?

These excerpts indicate that in the language of the teacher in the classroom and the language of the textbook writer the active voice is mostly absent when theory statements are made. The conclusion or result is portrayed as emerging from the external world of data. This pattern of rejection of personal terms and active voice centres the source of action in the scientific method. A reality is postulated outside the observer. The style of this language reinforces this externality and provides the basic epistemological tenet: by using the same method anyone and everyone will be lead to draw the same conclusion.

But this language can clearly be reformulated – to make the scientist the active agent interpreting the phenomena. For example, the texts quoted above from Mayfield *et al.* (pages 38, 41 and 44) could be rewritten as:

> The results shown in Figure 3.6 are interpreted by the statement that the acceleration A a is directly proportional to the unbalanced force F.
>
> We can make an hypothesis, based on calculations of this and of other similar collisions, that the action reaction forces are always equal in magnitude and opposite in direction.

Another key issue in the language mode that is a vehicle for positivism is what Gusfield calls "reductionism". He means by this term, the converting of 'fact' into 'imagery' or in other words the naming of experimental phenomena. It is the major way in which sense or relevance is attributed to scientific data. The language used in such a style leads towards one particular focus and away from others. This is particularly evident in this section of classroom dialogue:

No one has any idea of the two different types of electricity? Glen?

P: Charges.
T: No. They're both charges. Douglas?
P: Current electricity and –
T: Oh no. That's two different types of electricity. We're talking about two different types of charges. Yes?
P: Like charges and unlike charges.
T: Oh yes. Well that's right to a point but we give them different names. Yes?
P: Repulsive and attractive.
T: No.
P: AC and DC.
T: No, that's a different type of current. They're two (inaudible). No, No. Yes?
P: Static and current.
T: Does that give anyone a clue? Does that give anyone a clue? More than a clue. Andrew?
P: Positive and negative.
T: Right. Would anyone like to guess which charge is which: on this – which do you think would give you positive? This and this, or this and this? Come on. Like the old game – which hand? Which one would give me negative? Who says glass and

> silk? Who says this one? What do you think, Jenny? Do you think that would give me negative?
>
> (Grade 8 Science class)

It appears to me in this extract that the teacher's language conveys the meaning that the naming of theoretical constructs is outside our control. The dialogue seems to imply that there are clues to the naming of constructs in the experimental data itself.

A very common technique of starting a lesson or commencing a Chapter in a book is to set up some tension around some experimental phenomena. As the lesson or chapter unfolds the conflict is resolved by a change away from a commonsense perspective towards a new theoretical perspective. In this sense the flow of action in the lesson or chapter can be termed dramatic. Clearly the teacher or author intends to persuade the student of certain conclusions. But what Gusfield discovered was that the persuading is done by presenting the supposed "external" world. The language used in this dramatic presentation is emptied of feeling and emotion, it is descriptive but without imagery and therefore flat and "factual". The style of the communication portrays the new theoretical constructs as cognitive, logical and derivative rather than affective, creative and imposed. For example, consider extracts from the first seven paragraphs of Chapter 3: Force and Motion.

> ". . . From our everyday experience we seem to need force to keep things moving or to change their motion . . .
> . . . Our experiment suggests that motion at constant velocity requires no force. It could be argued that this is a generalization about all objects made from an experiment in which we studied only one object. However, there is nothing unique about the puck. We could put a banana on it and it would still continue with constant velocity if no force were applied. Since the same result is obtained no matter what object we use or how many times we repeat the experiment, we assume that it holds in general."
>
> (Mayfield *et al.*, 1973, p. 34–35)

This passage illustrates the lengths to which some authors will go in order to appear logical!

In conclusion, it appears that it is possible to recognize a language style which produces and/or reinforces a positivistic consciousness in the receivers of the communication. What is very significant also is that an alternative style which would convey the action of the human mind on scientific data, and portray science as a human construct is a relatively simple change.

2. Contingency Content Analysis

Content analysis is a research method for investigating any problem in which the pragmatics of communication serves as the basis of inference. The goal of content analysis is to present a systematic and objective description of the features of communication. An important characteristic of this method is that it is an "unobtrusive" research technique.

The particular type of content analysis used in this report is due to Osgood (1959), and is termed *contingency* content analysis. It is a research design that measures the relationship of two or more variables within a complete communication act, for example a lesson or a book. Osgood's work supports the following hypothesis: that inferences can be made as to the associative structure of a source (teacher, author) from the contingencies among terms in the communication act (lesson, book).

To apply this technique the communication is first divided into units according to some relevant criterion. A coder then notes for each unit the presence or absence of each content category for which he is coding. (For a complete list of the content categories see Appendix A). The contingencies or co-occurrences of categories in the same unit are then computed and tested for significance against the null (chance) hypothesis. And finally, patterns of greater-than or less-than-chance contingencies may be analysed. (The complete specifications of the data are contained in Appendix A).

When a complete list of significant contingencies which were greater-than-chance, were assembled the following two relevant clusters were constructed.

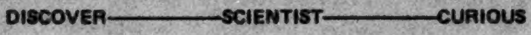


FIGURE 1

A cluster of significant contingencies

Thus the interpretation of the results appears to agree with the stated aim of the writers of the text (Heading *et al.*, 1972) which was to portray the scientist as a curious investigator of nature.

This second cluster of significant contingencies is much more relevant to the purposes of this article. Here the technique appears to uncover a web of connected categories that reveals a very positivistic view of the scientific endeavour and which again agrees with the stated intentions of the authors of the text.

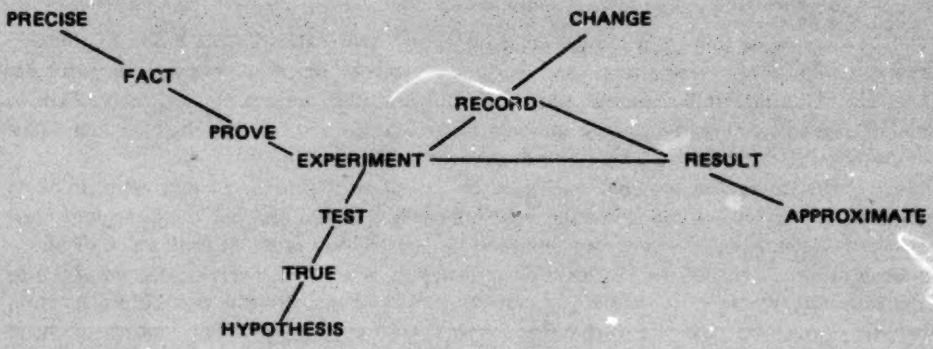


FIGURE 2

A cluster of significant contingencies

In summary, despite the limited use so far of this technique, it does seem to be capable of supporting the hypothesis that the language of some science texts do contain an embedded positivistic slant on knowledge. The application of this technique to the language of science lessons is presently in progress.

The predominance of the positivistic view

This final section offers a brief statement of two hypotheses to explain the persistence and predominance of positivism in the classroom. No doubt there are other hypotheses, but certainly it could be argued that they would be subsumed by the two global hypotheses mentioned below. The fist hypothesis is argued at greater length by Whitty (1974); and the second is derived from Lundgren (1977).

1. The Culture of Positivism

A substantial argument can be mounted to sustain the hypothesis that we live in a "culture of positivism" (Whitty, 1974). The work of Schutz (1962) on the common-sense experience of the world is the basis of the argument. He describes the public, common-sense mode of thinking about the world:

> By the operation of these constructs of common-sense thinking it is assumed that the sector of the world taken for granted by me is also taken for granted by you, my individual fellow-man, even more, that it is taken for granted by 'Us'. But this 'We' does not merely include you and me but 'everyone who is one of us', . . . Thus, the general thesis of reciprocal perspectives leads to the apprehension of objects and their aspects actually known by me and potentially known by you as everyone's knowledge. Such knowledge is conceived to be objective and anonymous i.e. detached from and independent of my and my fellow-man's definition of the situation, our unique biographical circumstances and the actual and potential purposes at hand involved therein.
>
> (Schutz, 1962, p. 12)

The key words here are obviously "objective" and "anonymous". This mode of thinking operates with a notion of natural, good and true knowledge which is detached from particular knowers. A culture that embodies such common-sense thinking can be labelled positivist in that it displays no interest in the grounds of knowledge and assumes that the knowledge produced by such thinking is simply true knowledge of Nature.

The Schutzian analysis then goes on to claim that this common-sense thinking takes account of the social distribution of knowledge. It assumes that there are individual differences in the actual knowledge possessed by people. Any single person's stock of knowledge is at any particular stage of his life considered to have zones of varying degrees of clarity and precision. In other words this common-sense thinking assumes that there is a distinction between those who possess expert public knowledge in certain zones (e.g. science teachers) and those who lack this knowledge (e.g. pupils). As Schutz succinctly states it: "The knowledge of these individual differences is itself an element of common-sense experience".

This common-sense understanding of the social distribution of knowledge leads directly to a transmission model of science teaching. It is assumed that the science lesson is the coming together of one who has knowledge with many who don't have knowledge. The culture interprets school science as simply the remedying of the inexactitudes, lacunae and subjective aspects of the common-sense knowledge of the pupils. The prevailing perception of school science can be described as the uncritical dissemination of scientific knowledge by the teacher to the pupils. This process tends as Popkewitz (1972) suggests to be simply "the replacement of one sort of comforting mystification with another".

Granted the validity of this Schutzian analysis it then follows that the culture of the classroom can be characterized as essentially positivistic. The embeddedness of this stance in the common-sense thinking of the actors in the classroom creates enormous difficulties for anyone trying to change this conception of the teaching-learning process. The new sociology of education (Young, 1971) provides a plausible critique of this mode of perceiving knowledge in the classroom. But what these critics wish to transcend – in their terms a 'false' consciousness – is arguably much more pervasive than they grant; and the task of changing this consciousness much more complex than they concede. Therefore, I am inclined to agree with the pessimistic conclusion of Whitty (1974) that it has proved an impractical task to successfully challenge the positivistic culture of the classroom with the new sociology of education as heralded in Young (1971).

2. The Social Constraints of the Classroom

The teacher's role in the classroom can be divided for present purposes into two parts – the control of knowledge and the control of the social order. These can be distinguished for analytic purposes, but clearly much of what the teacher does in a classroom is concerned with both of these roles simultaneously. The aim of this section is to look closely at the interaction of these two roles of the teacher. The argument will seek to establish that the teacher's role in maintaining the social order leads to a reified conception of knowledge and a transmission model of teaching.

The starting point for this discussion will be the communication patterns in the classroom. From a total of ten Grade 8 science lessons the teacher-student discourse was taped, transcribed and then classified according to a scheme developed by Sinclair and Coulthard (1975). This type of reduction of classroom discourse to general categories has been extensively pursued by Lundgren (1977); and a comparison of his classification scheme and that of Sinclair and Coulthard is contained in Appendix B. The results of the present study and a sample of Lundgren's results are presented in Table 1.

TABLE 1

Categories of classroom discourse

	*Lundgren study**		*Present study+*	
	Teacher	*Pupil*	*Teacher*	*Pupil*
Focussing	88%	12%	100%	0%
Opening	86%	14%	97%	3%
Answering	12%	88%	3%	97%
Follow up	81%	19%	100%	0%

*Lundgren sample : 17–18 year old Mathematics pupils.
+Present sample : 13–14 year old Science pupils.

These results reveal striking regularities in teaching. The data shows that often teaching is a monotonous process where the teacher initiates, structures and solicits, the pupils answer then the teacher reacts to the answer. The teacher dominates the communication and the pupils who seldom take the initiative, are restricted to short sentences and obvious answers. Lundgren found a remarkably similar pattern of communication vertically up and down the school grades, and horizontally across school subjects. This pattern has been termed the 'recitation' pattern and Lundgren hypothesized that it is functional in relation to the classroom situation. The explanation for its existence is presumed to be in the way in which the classroom contexts governs the actors. On the one hand the teacher appears to be in control of the process, but the regularities must mean that he himself is strongly under the control of other factors. Thus the first point to be made here is that these factors restrict the language patterns of the classroom, which in turn restrict the pupils' opportunities to broaden their perspectives; and they give rise to a transmission model of unexamined knowledge.

The recitation pattern as a teaching-learning strategy has received trenchant criticism (Hoetker & Ahlbrandt, 1969). According to Lundgren its endurance is explained in part by the necessity of the teacher to perform the *two* roles of control – the control of knowledge and the social order.

> "The recitation pattern is the only communication pattern which gives the illusion of instruction and at the same time guarantees a certain amount of social control."
>
> (Lundgren, 1977, p. 202)

The subtlety of this pattern has been explored by Lundgren and he has identified a process he terms 'piloting', which is a not uncommon pattern used by teachers. 'Piloting' refers to the strategy whereby the teacher simplifies topics so that pupils can answer a simple chain of questions, usually by a process of reading internal clues, thereby giving the impression of understanding. This success is achieved despite the fact that they could not possibly understand the basic principles. According to Lundgren, 'piloting' is explained by the teacher's transmission model of learning and its functional response to the social conditions of the classroom.

The recitation pattern is also partly explained by the demands of the curriculum: the pressure to 'get through the list of topics'. Under this type of applied stress it is clear that the communication patterns used by a harassed teacher will be such that the process of teaching can proceed with the minimum amount of interruption. This mode of discourse avoids the risk of problems arising and needing to be explored. There is just not enough time to correct all the gaps in the learning of thirty young people.

In conclusion, it seems at this stage, that classroom teachers to a large extent cannot avoid the positivistic role that the social constraints impose on them. New perspectives may offer teachers explanations and interpretations of what they do in the immediacy of the classroom interaction. But the likelihood of a change occurring is rather remote.

References

BARNES, D. *From communication to curriculum.* London: Penguin, 1976.

BELLACK, A.A., KLIEBARD, H., HYMAN, R. & SMITH, F.*The language of the classroom.* New York: Teachers College Press, Columbia University, 1966.

BERNSTEIN, B. A sociolinguistic approach to socialization; with some reference to educability. In J.J. Gumperz and D. Hymes (eds). *Directions in sociolinguistics.* New York: Holt, Rinehart & Winston Inc., 1972, 466.

FEYERABEND, P.K. Against method: Outline of an anarchistic theory of knowledge. In *Minnesota studies in the philosophy of science*, (Vol. 4). Minneapolis: University of Minnesota Press, 1970.

FLANDERS, N.A. *Analysing teaching behaviour.* Reading: Addison & Wesley Publishing Co., 1970.

GORBUTT, D. The new sociology of education. *Education for Teaching*, 1972, 89, 19–27.

GUSFIELD, J. The literary rhetoric of science: Comedy and pathos in drinking driver research. *American Sociological Review*, 41 (1), 16–34.

HEADING, K., PROVIS, D., SCOTT, T., SMITH, J. & SMITH, R. *Science for secondary schools* (Vol. 1). 3rd ed. Sydney: Rigby Ltd., 1972.

HOETKER, J. & AHLBRANDT, P.A. The persistence of recitation. *American Educational Research Journal*, 1969, 6 (2), 145–167.

HOOKER, C.A. Philosophy and meta-philosophy of science: Empiricism, Popperianism and realism. *Synthese*. 1975, 32, 177–231.

KUHN, T. *The Structure of scientific revolutions*. (2nd ed.). Chicago: University of Chicago Press, 1970.

LUNDGREN, U.P. *Frame factors and the teaching process*. Stockholm: Almquist & Wiksell, 1972.

LUNDGREN, U.P. *Model analysis of pedagogical processes*. Lund: Liber Publishing Co., 1977.

MAYFIELD, J.M., PARHAM, R. & WEBBER, B. *Fundamentals of senior physics*. Book one. (2nd ed.). Melbourne: Heinemann Educational, 1973.

MILLS, C.W. Language, logic and children. In I. Horowitz (ed.). *Power, politics and people*. New York: Oxford University Press, 1963.

MUNBY, A.H. Some implications of language in science education. *Science Education*, 1976, 60 (1), 115–124.

OSGOOD, C.E. The representational model and relevant research methods. In I. De S. Pool, *Trends in content analysis*. Urbana: University of Illinois Press, 1959.

POPKEWITZ, T.S. The craft of study, structure and schooling. *Teachers College Record*, 1972, 74 (2), 155–165.

SCHUTZ, A. Common-sense and scientific interpretation of human action. In M. Natanson (ed.), *Alfred Schutz: Collected papers* Vol. 1. The Hague: Martinus Nijhoff, 1962, 12.

SINCLAIR, J. Mc H. & COULTHARD, R.M. *Towards an analysis of discourse*. London: Oxford University Press, 1975.

TOULMIN, S. *Human understanding*, Vol. 1. Oxford: Oxford University Press, 1972.

WHITTY, G. Sociology and the problem of radical educational change. In M. Flude & J. Ahier (eds), *Educability, schools and ideology*. London: Croom Helm, 1974.

YOUNG, M.F. (ed.). *Knowledge and control*. London: Collier-Macmillan, 1971.

YOUNG, M.F. School science – Innovations or alienation? In P. Woods & M. Hammersely (eds). *School experience*. London: Croom Helm, 1977.

Appendix A

Contingency content analysis of *Science for secondary schools* Vol 1, by Heading *et al.*, 1972.

*CATEGORIES**		*SIGNIFICANT CONTINGENCIES*	
		.01 level	*.05 level*
experiment	(47)	prove	record, results
learn	(39)	science	
prove	(28)		infer
explain	(25)		scientific method, curious
scientist	(25)	problem	curious
change	(23)	record	
cause	(16)		predict, science
study	(14)		science
results	(13)		approximate
inquire	(13)	curious	
test	(11)	experiment, tentative	true
assume	(11)	record, approximate	
note	(10)	hypothesis	
hypothesis	(9.6)	scientific method	true
fact	(9)		precise, prove, tentative, curious
true	(8.4)	approximate	assume
record	(8.4)	results	inquire
precise	(8.4)		scientific method, curious
understand	(7.2)	science	
discover	(6.6)	chance, assume, knowledge	scientist, precise, scientific method
curious	(6)	effect	science
knowledge	(4.8)	science	
evidence	(4.2)		inquire
theory	(3.6)	assume	explain, understand, inquire
tentative	(3.6)		infer
invent	(3)	create, discover, control	
problem	(2.4)	precise, scientific method	tentative, inquire
create	(1.8)	discover, control	science
predict	(1.8)	science	evidence, knowledge
logical	(1.2)	precise, assume	
chance	(1.2)	scientific method, tentative	precise
scientific method	(1.2)	tentative	

*The number listed with each category represents the percentage of the 166 units into which the text was divided, which contained that category. For example, the category 'test' appeared in 11% of the 166 units.

Appendix B

Classifications of Classroom Discourse

Lundgren's pedagogical moves (1977)	Sinclair's system of analysis (1975)
Structuring (STR) – which set the context for subsequent behaviour.	Focusing (FOC) which is to talk about the discourse, to tell the children what is going to happen or has happened. The teacher stands for a moment outside the discourse.
Soliciting (SOL) – moves which are intended (a) an active verbal response (b) a cognitive response e.g. attending (c) a physical response.	Opening (OP) – moves which cause others to participate in an exchange. These moves are in three categories (a) inform (b) direct (c) elicit.
Responding (RES) – moves which reciprocate soliciting moves.	Answering (AN) which is a response appropriate to the category of opening move.
Reacting (REA) – moves which are related directly to the three other moves but need not be initiated by them.	Follow up (FU) – which is to let the pupil know how well he/she has performed.

Research in Science Education, 1978, 8, 25-33

THE ROUGH THEATRE OF SCIENCE TEACHING IN THE COMPULSORY YEARS
RATIONALITY & ARTISTRY

Rod Fawns

Graves crystallizes for me the improbability that current science teaching can be an effective instrument of general education:

IN BROKEN IMAGES

He is quick, thinking in clear images;
I am slow, thinking in broken images.

He becomes dull, trusting to his clear images;
I become sharp, mistrusting my broken images.

Trusting his images, he assumes their relevance;
Mistrusting my images, I question their relevance.

Assuming their relevance, he assumes the fact;
Questioning their relevance, I question the fact.

When the fact fails him, he questions his senses;
When the fact fails me, I approve my senses.

He continues quick and dull in his clear images;
I continue slow and sharp in my broken images.

He in a new confusion of his understanding;
I in a new understanding of my confusion.

Robert Graves

Before our compulsory science classes we often appear sharpest in the understandings of our broken images and dullest in the confusions of our clear images.

When we look about us for an education whose impact is immediate, for a comprehensive and penetrating pleasure, we see one of the major opportunities and responsibilities of science education in the compulsory years to be effective in teaching "about" science as well as "in" science – science as an active humanity concerned with basic social values.

The problem is to resolve the conflict between the existence and the essence.

Meaning and relevance we readily concede to be more than just a property of the "new course content" or recommended "approach", and we regularly express concern that the dynamics of the representation are mistaken for the dynamics of the matter to be represented. However, little attention seems to be given in teacher education or research to the developing personal critique of science teaching performance which encompasses social, aesthetic as well as rational values.[1] Like perception, the quality of the performance is felt to be inaccessible, related to previous experience, guided by only partly specifiable clues towards the achievement of a coherence sensed by each of us and only communicated in broken images.[2] In any case, doesn't the artist always rise above the prescriptions of the theorist?

Perhaps, but the use of theory is not restricted to prescription or prediction and it is likely that there is other human experience expressed in non-scientific theories which may at least structure or clarify the problem. Such theories we could predict to be testable against experience but certainly less objectively.

Science teaching which engages compulsory classes directly, often seems to me to bear an affinity to the tradition of popular theatre. This theatre takes many forms but there is one factor that is common - a roughness. Salt, sweat, noise, smell: the theatre doesn't pick and choose, the audience has to be the cast, the teacher more the director-producer.

If the audience is restive, then it may be more important to confront the trouble makers, or improvise a joke - than to try to preserve the unity of style in the scene. In the luxury of high-class theatre (the form 6 prac. and discussion) everything can be of a piece. In a rough theatre any offering must be used, developed, recycled and reviewed at another level. The arsenal is limitless: the aside, the topical reference, the local jokes, the exploiting of accidents, the tempo, the noise, the relying on contrasts, the shorthand of exaggeration. This popular theatre, freed of unity of style actually speaks a very sophisticated and stylish language. The class seems, often inexplicably, to have no difficulty in accepting inconsistencies of language, or in darting between realism and suggestion, fact and metaphor.

Here is part of a live script from the rough theatre of a form 3E ("lowest ability") science class investigating nothing less than 'The Nature of Light'.[3]

We think there is a dramatist at work here.

(Apparatus which includes a light globe and a razor blade occupies the centre of the room)

T	What is this this thing you're calling the light? What is it made of? Is it made up of - is it made of liquid, stuff like water and runs around the room ---	Spontaneous use of a classroom idea - notice how it is met to expose its weakness.
S	Nah! It's solid, it's solid.	
T	It's a solid. Uh? It shines and hits you in the face and you feel it do you?	
S	No! (chorus) It's a gas! It's like a white powder.	
T	It's a sort of gas, what is a gas made of?	
Davis	Particles.	
T	Particles, what are they called - anyone know? What are gases --- come on girls --- made of?	
(S	atoms)	
T	These particles in a gas. Who has (S atoms, atoms) seen some gas?	Levels of Analysis by Teacher - complex seeing - concrete resolution.
S	I have.	
S	No-one has seen gas.	
T	You blow against your hand. You can feel that gas. It's made of um! --	
S	Air atoms.	
S	Hydrogen atoms - gas - atoms Sir! Sir!	
S	You cannot see gas unless you're right in it.	
T	Noble.	
Noble	Sir, why can't you see gas?	
T	You can see some sorts of gases	
S	See!	
T	Right! Hands down.	Humour or sarcasm? - at this point it is hard to tell - but the reflex, the level of response is unexpected - paradoxical surprise - . . . easy? . . . perhaps.
Noble	Sir, can I draw the curtains?	
T	Listen, listen that's all you come here for Noble is to draw the curtains. (Laughter)	

T (with measured slowness)	We'll do a special lesson for you Noble you can come and sit on your own and we'll draw the curtains for you – that will really suit you well – drawing the curtains. Noble, we will draw the curtains but we'll do something more important once we've done it. Now listen – let's go back in time again those 300 years.	the seriousness of the venture emphasized without confrontation. Noble moves next to the teacher who senses the audience is with him and breaks to establish the illusion.
Noble	Sir, are we going to do something with that chart on light bulbs over there on the wall?	
T	Yes, something on light bulbs. Do you want to go over and read them? O.K.! Off you go (Noble goes over to the chart).	Noble breaks back – teacher attends using the whole room as a stage.
T (almost confidentially)	Something about 300 years ago --- To yourself please Noble --- 300 years ago these early scientists didn't have much apparatus but Gino said they 'had good brains'. All they did, they didn't do many experiments but they used to make --	A theatre of narrative – no precipices of disbelief to scale – gentle reasonableness. Notice how earlier insights are recycled – progress as in curves – "good brains" – enhances the narrative. Engaged responses?
S	Materials ------ guesses,	
T	Guesses! Yes guesses what's the word for these scientific guesses, not estimates --- begins with "Th".	
S	Theories! Theories!	
T	Great! They used to make these theories. Years ago talking about light there were two great men, great people. One of them very famous whose first name . was Issac.	
S	Isaac Newton.	The illusion breaker is not confronted – the interjection is turned into an almost serious question. Even the firm response is respectful . . . . 'you know it' . . . .
T	Isaac Newton. Great. Isaac Newton.	
Noble	He made the light bulb.	
T (reflective)	Did he make light bulbs? We'll see about this.	
Noble	It says up there ---	
T (firmly)	I don't want to know what it says – you read it you know it ---	
T	Isaac Newton is one of those people like Davis over here who thinks light is made of ---	recycles Davis' hunch but new, now it is a 'movement'!
S	Particles.	
T	There was another man who lived at roughly the same time	a clever chip – low ability!! funny enough to break the illusion – the teacher now has to move quickly to maintain the sense of the narrative. Look at the question! "Who believes . . . .?" High risk? And look who offers . . . enthusiasm-focused. Note the essential repetition to associate waves with . . Huygens . . mainly for Noble.
S	Magnus Pike (Laughter)	
T	No it wasn't Magnus Pike. He was a very important scientist in his day. His name was Christian Huygens. He believed that light wasn't particles but rays like waves. Who believes that light is waves? Who wants to be Christian Huygens?	
Noble	I'll be him. I'll be him.	
T	Hands down. Gino you've had a lot to say and you've said most of it. Noble you think it is waves. Do you want to be Huygens.	
Noble	Yes, Yes!	Noble is placed on the 'stage' (he actually wasn't sitting) Can Noble do it . . . or be led to it? Everyone knows its a risk . . . . AT least it's a statement – small beginnings.
T	Right stand up Noble. Now there were often meetings between the great men and Old Isaac would come along and say "Christian you know light is really particles, your thinking is a load of rubbish". What do you think Huygens would say?	
Noble	He would say, "No it isn't ---- it's made of waves".	
T	Good and you'd say, well I think it's made of particles because --- This is your system.	

	Particles bounce around the room and particles bounce around room – go on tell him.	
Davis	Particles bounce around the room.	
T	That's his theory – why do you think it is waves?	
Noble	Because it flows around all the room. If it was particles it would bounce to one wall would be dark and the other end might be light.	
T	Oh!	
Noble's mate	If it was waves, flow around the room it will be light all at the same time and it will get to the corners as well.	Noble does it! – a magic moment? The teacher is in the audience. Davis defends well but Noble has a bigger question. Should the teacher have explained the fallacy in Noble's argument?
Davis	No! No!	
T	Good point, good point.	
Davis	Light *doesn't* get to corners.	
Noble	Sir – what would happen if the particles bounding around the room bumped into one another?	
T	What would they do then?	
Noble's mate	Nothing just sort of . . .	At this point the teacher and class are lost in the image of colliding particles of light. Look at the questions being asked by the class! Low ability??
T & Noble (speculatively)	Explode! Explode into little pieces?	
Noble's friend	Sir, Sir. He says light's a particle but just say it goes to one corner of the room. What happens then?	
T	Only one particle?	
S	There may be lots.	
T	There may be millions of particles.	
S	Sir, how long would it bounce around?	
T (now speaking vigorously)	I don't know - - - - - - Well done! You boys great listen we'll now go on to an experiment. Well done Noble, well done Davis and our two assistants. A good chat it was. Now my turn to talk – these great men had their followers – these two and others were particle men perhaps these two were wave men.	Moving on, while the excitement is there. "Well done . . . like a football coach, appropriately understated. 'A good chat it was'.
T	If nothing else had been said there would still be these two groups of people today. There are not. Something happened. Someone proved his point.	He doesn't tell . . . he hasn't at any stage. Should he? High engagement as they approach the 'experimental . . . . . proof' and for this also the teacher has props (simple apparatus in the centre of the room). The girls have said nothing – how do you feel about that?
S	Whose?	
T	I'm not telling.	
S	I know, it's the wave man's theory.	
T	Listen carefully and we'll see.	

Notice how script-like the transcript is:

1. There are no long teacher monologues and an attempt is made to give members of the class significant spaces. The teacher is a narrator to the lesson the class is shaping.

2. There is simple, direct engagement between the teacher, the class and the issues. There is one play – not two. The teacher-producer and the class are in the same theatre if not on the same stage! Fantasy and invention are legitimate methods. What is important is that reality be correctly understood and accurately represented and in this area that is a subjective judgement.

3. The whole narrative is convincing because it is complex enough – language stylish yet plain – stylish because it is plain, yet it works at several levels. It has certain epic qualities. It draws its dramatic effect as much from the structure of the content (the relationship between basic ideas/facts/principles) as from the structure of the enactment (scenes, pace, level). The two are separable in theory but rarely in practice here.

4. The coherence of the narrative cannot therefore be summarized in terms alone of externally imposed structures or the quality of the relationships between the teacher and individuals in the class. It is also intimately related to "what" questions around what is worth teaching – questions of economy and power? There are scenes clearly discernible and meaningful but also powerfully related.

5. The relevance, meaning of the narrative, is that attributed by all participants, individually and collectively. Scientific 'literacy' is certainly being offered but not in any narrow sense for reasons of higher productivity. Relevance can be seen perhaps as a quality of the performance.

6. There is great personal respect and commitment to the learning and enjoyment of the class. It is young teaching – a very high level of intellectual and emotional energy is invested in the lesson. He says what he has to say to particular people, in particular situations, to reach a particular end for particular reasons He. sees the dialectic as potentially dramatic and not just valid.

This analysis of the lesson was made in the partial light of clues or distinctions proper to the dramatic theory of Bertolt Brecht.

Brecht's importance to modern theatre is accepted, Demetz (1962), Gray (1976), Williams (1968), Brook (1977), indeed held to be "conclusive" by Williams in a particular dramatic world – that of the individual against society – represented in three or four great plays – Mother Courage, The Good Woman of Setzuan and the Life of Galileo and The Caucasian Chalk Circle.

What is this theatre's intent? First of all it is there unashamedly to provide happiness – a "theatre of delight' and any teacher that can give delight has earned her place. Along with serious, committed and probing work there must be irresponsibility. But fun for fun's sake is seldom enough. Brecht's unique aim was to create a precise effect on an audience for whom he had total respect.

The effect he called 'alienation' – an estrangement from what is being observed in order to see it anew but differently to establish more complex seeing. In analysing and separating the object into individual parts, he believed the thinker produces at first merely dissociated elements but he has set a dialectical process in motion and the true recognition of an object hitherto mistakenly seen is only reached when the negation is negated. Brecht's great success as a craftsman of the theatre springs from his ability to represent complex processes and actions in a single memorable "object" or "gesture". "Truth is Concrete" was a sign he had placed over his work table for many years. Brecht's intention was not merely to make the familiar unfamiliar, but to lead on to a fresh vision of reality in a more real sense. By suspending belief he hoped that the illusion itself would become analyzable.

Alienation, central to Brecht's dramatic theory, also embodies many elements of social criticism – it is apparent in the fact that everything is something different from itself, that my activity is something other than mine. Brecht's theory of the "alienation of alienation" is thus related to, but in other respects clearly distinguishable from the theory of alienation described by Freud and Marx (Israel, 1971).

After long experiment, Brecht created a dramatic form in which men are shown in the process of producing themselves and their situation. The pure naturalist form of theatre, he felt, depended on a single materialist view, in which man discovers the truth about himself by discovering his real environment: the literal presentation of this environment is then a means to human truth. He felt there was drama in showing men in the process of producing themselves and their situations as opposed to discovering themselves in a given situation.

Clearly it's a theatre of consciousness. "Since the audience are not invited to throw themselves into the lesson as into a river, in order to let themselves be tossed indeterminately back and forth, the individual events must be tied together in such a way that the knots are strikingly noticeable. The events will not follow upon one another imperceptibly, but rather they must be able to pass judgement in the midst of them". The parts of the representation, therefore, are to be carefully set off against one another by giving them their own structure, that of lessons within the lesson (Brecht, 1964).

Such teaching would seldom seem linear nor would it mysteriously grow, its progress would be more in curves, and sudden leaps. The classes would give the illusion of being slower paced, with room to move; they would be reflective given time to talk and compare. The drama is drawn from the activity (including investigations) but mainly from within the narrative of the lesson.

Brecht's theatre proceeds by distancing and demonstrating. What he seized upon was the exclusion of all direct commentary, alternative consciousness, alternative points of view in fact of argument in the narrative in much of contemporary theatre.

Brecht's Life of Galileo (1974) is intended as narrative – rough theatre. As in previous lesson, different levels of meaning are interwoven. In Scene 3 Galileo's momentous act of scientific investigation – is isolated, interrupted, alienated by his friend's concern and the paradox of Galileo's (and science's) naivety is seen differently – a more complex seeing.

GALILEO: It is proved. The fourth can only have gone behind Jupiter, where it cannot be seen. There you have a star round which another revolves.

SACREDO: But the crystal sphere to which Jupiter is attached?

GALILEO: Yes, where is it now? How can Jupiter be attached to anything when other stars circle round it? There is no framework in Heaven, there is no fixity in the universe. There is another sun!

SAGREDO: Calm yourself. You think too quickly.

GALILEO: Quickly! Rouse yourself, man! What you have seen, no one has seen before. – They were right.

SAGREDO: Who! The Copernicans?

GALILEO: And the others! The whole world was against them, and they were right. This is something for Andrea! *Beside himself with excitement, he runs to the door and shouts:* Signora Sarti! Signora Sarti!

SAGREDO: Galileo, calm yourself!

GALILEO: Sagredo, excite yourself! Signora Sarti!

SAGREDO *turns the telescope away:* Will you stop roaring around like a lunatic?

GALILEO: And will you stop standing there like a cod-fish – when the truth has been discovered.

SAGREDO: I am not standing like a cod-fish, but I tremble lest it may in fact be the truth.

GALILEO: What?

SAGREDO: Have you entirely lost your senses? Do you really no longer know what you are involved in, if what you see there is true? And you go shouting about for all the world to hear: that the earth is a star and not the centre of the universe.

GALILEO: Yes! And that the whole, vast universe with all its stars does not revolve round our tiny earth – as must be obvious to everyone.

SAGREDO: So that there are only stars there! – And where then is God?

GALILEO: What do you mean?

SAGREDO: God! Where is God?

GALILEO: *angrily:* Not there! Any more than he could be found on earth, if there were beings up there and they were to seek him here!

SAGREDO: Then where is God?

GALILEO: Am I a theologian? I'm a mathematician.

SAGREDO: First and foremost, you are a man. And I ask you, where is God in your universe?

GALILEO: In us or nowhere.
SAGREDO: *shouting*: As the heretic Giordano said?
GALILEO: As the heretic Giordano said.
SAGREDO: That was why he was burnt! Not ten years ago!
GALILEO: Because he could prove nothing. Because he only stated it –

(*The Life of Galileo*, Scene 3, p 40-41)

A clue structure to a critique of teaching about science

Perhaps we can review and clarify Brecht's position (although this can really only be done by studying the plays and the critics) by contrasting the purpose and dynamics of his epic-narrative theatre with the dramatic-active theatre which he attacked.

Active Theatre	Narrative Theatre
The Key meaning is in the action – the movements and interactions of the class – often naturalistically portrayed	The Key meaning is in the narrative – what the teacher and class say and how they describe themselves and the situation – often as historical-social representation
Involves the student emotionally in the stage action	Makes the student an observer, but
consumes his capacity to act by	awakens his capacity to initiate by
appealing to empathy	demanding decisions from him
offers experience	offers a view of the world
suggestion	argument
the student is drawn into something	he is confronted with something
the student stands inside, experiences with others	he confronts and studies what he sees – thinks above the flow
feelings are preserved	feelings driven into becoming realization
Man is assumed to be predictable if not known	Man is an object of investigation
Man is unalterable	Man is alterable and altering
Suspense in awaiting the outcome of the interactions	Suspense at the process
one scene exists for another	each scene for itself
growth	montage
linear progress	in curves
evolutionary inevitability	sudden leaps
thought determines being	social being determines thought
dramatic power is drawn from the dynamics of the theatrical representation and the actions performed	dramatic power drawn from the debate – from the dynamics of representation and matter represented
Exciting, full of action and tensions	Slow paced, reflective, giving time to reflect and compare

This is clearly over schematic and expresses more perhaps differences in emphasis. Certain questions now need examination: Are all distinctions meaningful? Is there any coherence in these two forms? Is one closer, (clue structure used as a sort of checklist) to the style or values of current compulsory science teaching under the influence of epigenetic theory, which confers on psychological fact a pseudo-independence for social facts and confers on social facts the appearance of being mere expressions of individual personality (Schwab, 1969).[4a]

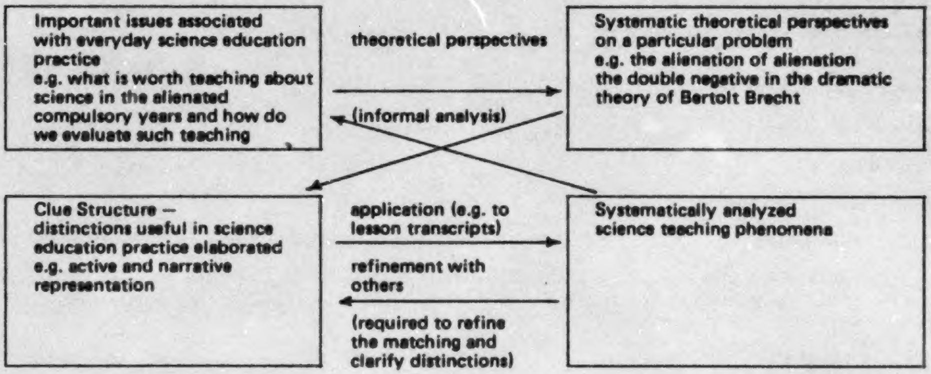
Such a contrast may be immediately useful in offering one critical framework against which a critique of current approaches to social representation of science — games, simulations, roleplays or historical case studies may be developed.

I believe Layton (1973) is right, science education like the classics, must re-establish itself as an effective instrument of general education by widening its role to include the classroom analysis of the external relations of science, the science-society interface will require a different self-consciousness associated with new theoretical support.[4] However we are bound to experience in the current economic climate continuing attempts to render science education accountable to only a limited number of outcomes. The questions of who is the audience of science education and who speaks for us remain.

The Science Masters' Association as far back as 1961 were urging in support of curriculum reform that "Science should be recognized and taught as a major human activity which explores the realm of human experience, maps it methodically but also imaginatively, and by disciplined speculation, creates a coherent system of knowledge. As a human quest for Truth it is much more subjectively human than is often realised. Science is concerned with basic values and is an active humanity." (Science Masters' Association, 1961). The bid for science as culture has only really just begun, and meanwhile the question of how we value and how we represent this age remains but hopefully less resistable.

Is there a research methodology implied here?

Not a methodology but perhaps an approach to exploring the value of a different theoretical perspective. It is represented in high gloss schematically below:


Important issues associated with everyday science education practice e.g. what is worth teaching about science in the alienated compulsory years and how do we evaluate such teaching
theoretical perspectives
(informal analysis)
Systematic theoretical perspectives on a particular problem e.g. the alienation of alienation the double negative in the dramatic theory of Bertolt Brecht
Clue Structure — distinctions useful in science education practice elaborated e.g. active and narrative representation
application (e.g. to lesson transcripts)
refinement with others
(required to refine the matching and clarify distinctions)
Systematically analyzed science teaching phenomena

References

BRECHT, Bertolt. *The life of Galileo*, Eyre Methuen, 1974.

BRECHT, Bertolt. A short organum for the theatre (1948). In Willett (ed.), *Brecht on theatre, The development of an aesthetic.* Methuen, 1964.

BROOK, P. *The empty space.* Pelican, 1977.

DEMETZ, P. (ed.). *Brecht. A collection of critical essays.* New Jersey: Prentice-Hall, 1962.

GRAY, R. *Brecht the dramatist.* Cambridge University Press, 1976.

ISRAEL, J. *Alienation from Marx to modern sociology.* Allen & Bacon, 1971.

LAYTON, D. *Science for the people: The origins of the School Science Curriculum in England.* London: Allen & Unwin, 1973.

MORPHET, C. *Galileo and the Copernican astronomy.* S.I.S.C.O.N., Butterworth, 1977.

SCHWAB, J.J. The practical: A language for curriculum. *School Review*, 1969, 78 (1).

SCIENCE MASTERS' ASSOCIATION AND THE ASSOCIATION OF WOMEN SCIENCE TEACHERS. *Science education: A policy statement.* London: John Murray, 1961.

WILLIAMS, R. *Drama from Ibsen to Brecht.* Pelican, 1968.

Footnotes

1. One experimental approach to the tacit-reflexive (personal theory building) dimension of pre-science science teacher education is presented in a paper presented to S.P.A.T.E., 1978, 'Theory out of Practice – a direct approach to Curriculum Studies', R.A. Fawns.
2. Science Teacher Education Project (S.T.E.P.) and Australian Science Teacher Education Project Materials (A.S.T.E.P.) provide useful external theory in the 'Art of Science Teaching' but where is the 'Performing Art'?
3. The lesson (part only) was recorded at Heathland Comprehensive School in London's western suburbs near Heathrow Airport. The young Physics teacher is in his third year of teaching. The tape is available from Science Teachers Association of Victoria Services at $5.00 per copy.
4. (a) Look to Vygotsky's 'Thought and Language' for a break from the Piagetian nexus, p. 20–24, M.I.T. Press, 1962.
   (b) Look for example to the Dramatic Theory of Bertolt Brecht variously reported in Williams and Brook (see references) and in other places.
   (c) Look for example at the recent articles in *Science/New Scientist.*
      - – Marwick, R. Is science a form of witchcraft? *New Scientist.* 5 September, 1974.
      - – Brush, S.G. Should the history of science be rated X? *Science*, 1974, 183, 1164–1172.
      - – Haberer, J. Politicalization in science. *Science*, 1972, 178, 712–724.
      - – Uri Geller & Science. *New Scientist*, 17 October, 1974.

Research in Science Education, 1978, 8, 39-43

SOME IMPLICATIONS OF THE INTRODUCTION OF ENQUIRY METHODS TO THE INTEGRATED SCIENCE CURRICULUM IN INDONESIA

G. Wayan Seregeg

Introduction

Indonesia had been under Dutch control for about 350 years before proclaiming independence on 17 August 1945. The present educational system stems from the Dutch colonial administration and very little change has occurred in the 30 years since independence.

The Dutch colonial administrators in Indonesia aimed at the workers who could read, write and count. Therefore, the aims of Dutch education then were to teach the three R's, i.e.: writing, reading, and counting. A system of *sekolah desa* (village school) was introduced with classes ranging from grades one to three. The media of instruction at the school were local languages. The school practices were very rigid since the Dutch government wanted strictly disciplined workers. Teachers were held in very high esteem and whatever the teacher said was accepted without question.

Suwardi Surjaningrat, also known as Ki Hadjar Dewantoro[1], a man of princely birth, developed the conviction that the colonial type of education was contrary to the Indonesian Culture and philosophical concepts and that it did not fit Indonesian conditions. Therefore, in 1922 he established the *Taman Siswa* schools (pupils garden). His basic idea was to have a definite foundation of culture and thought. The line of national culture needed to be adapted to new circumstances and necessities, but should never be broken. His idea of having a new system of education did not come from contempt for western virtues and ideas. Rather it was built on the idea of a well-developed system which was a blending of both occidental and oriental cultures. He said that a purely western education would produce maladjusted individuals who could neither fully digest western ideas nor at the same time feel quite at home in Indonesian surroundings.

The *Taman Siswa* schools rapidly grew in number and spread all over the West-Indies.[2] The influence of Dewontoro increased steadily through the years. But Mauldin (1961, p. 129–134) stated that in spite of this, the *Taman Siswa* philosophy, which kept the nationalistic spirit burning for many years, has had practically no influence on the present system of Indonesian education.

Radcliff (1971) claimed that the educational theory of Dewantoro was carefully developed to solve the problems posed by the collision of east and west, of the Indonesian and Dutch-European cultures. It was not a purely traditional, anti-colonial reaction; it was consciously progressive and open-minded, accepting from the west whatever was the value and a positive contribution to Indonesian life, but seeking to graft this on to a healthy and thriving indiginous stem.

The central principal of *Taman Siswa* was known as *sistim among* (among system). *Sistim among* means the construction of education around the child, making the maximum use of his own self-educational instinct; it was a child-centred education. Dewantoro stressed the importance of the child's culture as an important part of the whole culture of society. *Taman Siswa* theory stressed independence as a basic principle of education. It made an explicit connection between the intellectual independence of the child as a goal of education, and the implications of this for national independence. It had a particularly well-developed theoretical foundation (Radcliff, 1971).

---

[1] In 1959 his birthday, 2 May, was proclaimed as National Education Day.

[2] West-Indies was the name of the country that is now Indonesia.

Since 1965 many attempts have been made to revise the curriculum, but these revisions have been neither profound nor effectively implemented (Crellin, 1976). In 1973 the Ministry of Education and Culture designed the *Proyek Perintis Sekolah Pembangunan* (PPSP = Development School Project). Among the more important decisions made at the time were: (1) the teaching would be child-centred; and (2) the system would involve continuous progress assessment for each student.

The PPSP is technically conducted by the BP_3K (Office of Educational Research Development). The teaching method to be used was considered very carefully and it was decided to adopt a self-instructional, self-paced modular mode of instruction. In the Indonesian curriculum plan, "modular instruction" means teaching with the use of an instructional module.

Thomas (1974) defined "module" as a packet of suggestions for teachers and of learning materials for students that can be used for persuing specific learning objectives for a period of time that may be as short as fifteen minutes or as long as six or eight class periods distributed over a series of three or four weeks.

What then is the role of the teacher? He does not lecture, nor does he conduct question-answer sessions. Rather, he moves about the class, observing each student's progress and assists them in interpreting the printed materials, in reading the diagram and in dissecting the specimen the pamphlet requires him to analyze. This method of instruction then turns the teacher into a guide, motivator, supervisor, and evaluator.

In this study, the teaching method referred to above is taken to imply an enquiry approach even though the word "enquiry" is not used.

The enquiry approach is an innovation in Indonesia. It came from the American education system. Introducing an innovation to a country quite different in its socio-cultural background is a very difficult task. The findings of many studies emphasize that in a non-western culture, no western ideas can be taken for granted (Prince, 1969). Reference has been made to the fact that some western ideas fail to motivate non-western students. There is reason to believe that the cultural background of the students in western countries is different from those in non-western countries. The factor affecting motivation is rooted in the cultural background.

Lewis (1972) underlined the possibliities of applying science to the exploitation of the natural resources. He described that man can innovate, that the material world can be controlled and modified to man's advantage by man's own endeavours. The "mental mutation" that must be brought about is the belief that man can control his own destiny, that solutions can be found to problems, and that through the exercise of a proper spirit of enquiry it is possible for man to help himself. So, if education is to contribute to socioeconomic development, enquiring minds must be developed together with the confidence that problems can be solved and scientific habits of thought and action must be established.

What is the rationale of "enquiry"? There has been confusion over whether enquiry refers to a method of teaching, a method of learning, or something you learn. Bruner, the most eloquent defender of learning by discovery[3] said that enquiry is a method of learning.

---

[3] There are a number of terms used to describe the activity-oriented learning, i.e.: enquiry, discovery and and investigation. In this study, they will be used interchangeably.

In his article, "Art of Discovery", he claims four advantages of enquiry learning: (1) enquiry learning increases intellectual potency; (2) enquiry learning increases motivation; it strengthens the student's tendency to carry out his learning activities with the autonomy of self-reward or the reward of enquiry itself; (3) enquiry learning teaches the student the techniques of discovery; solving problems through enquiry develops a style of problem-solving that serves for any task one may encounter; (4) enquiry learning results in better retention of what is learned (De Cecco & Crawford, 1974, p. 354-363).

Bruner's descriptions of enquiry learning is typical of a western concept. The concept of enquiry places a higher social value on intellectual activity than dexterity. The problem is whether the concept will successfully motivate the Indonesian student to learn, to increase intellectual potency, and to strengthen the student's autonomy of self-reward. In that country, only a few grasp the relationship between cause and effect which can be studied and controlled. The student's views are rooted in traditional concepts of the nature of the world around them. The majority of children in the country develop the concepts of physical world based upon belief in the influence of supernatural powers. This teaching approach totally contradicts the accepted values of traditional society, in which the authority of the family head in the home and the teacher in school is accepted without question.

This study will examine the effectiveness of the enquiry approach applied in the primary school student in a non-western culture, Indonesia.

Design of this study

The problem of design is how to control the variables and how those variables are to be measured. Measurement does not provide information on all the qualities which characterize the object or event being measured. The major criticism of the data analysis was that many statistical analyses relied too heavily upon sophisticated statistics. Hudson stated that "sophisticated statistics poised upon techniques of rustic simplicity" (Cohen, 1973).

There are many variables which will be considered, as they influence the learning outcomes. Some of them are teaching strategies, student ages, sex balance in a class, socio-economic levels, parents education, parents occupation, school facilities, and language as a medium of instruction. They are so interwoven that in any given context several of them will be involved simultaneously.

Good and Hatt (1952) identified four problems which can be stated in four propositions, i.e.: (1) Human behaviour changes too much from one period to the next to permit scientific, exact predictions; (2) Human behaviour is too elusive, subtle, and complex to yield to the rigid categorization and artificial instruments of science; (3) Human behaviour can be studied only by other human observers, and these always distort fundamentally the facts being observed, so that there can be no objective procedures for achieving the truth; and (4) Human beings are the subject of such predictions and have the ability deliberately to upset any predictions we make. Also Kaplan (1971) stated one of the basic characteristics of human behaviour is that it is purposive, but that purposes, and their corresponding goals and values, are far from being as single and as simple as measurement seems to require.

The instrument used in this study was a non-verbal, multiple-choice test. It was adopted from the *Concepts in Science: Teaching Tests* (New York: Harcourt, Brace, Javanovich, 1976) and adapted to Indonesian conditions – fauna, flora, non-living things, the education system, and the language.

There are several methods of logical demonstration which could be employed. But this study will use only one, the cross-sectional design. This type of research is to find a measure of a certain variable and to compare this with observed patterns at the time of study.

Correlation between two variables will be analyzed by the method of concomitant variation. This method holds that if a change occurs in the amount of another variable in two or more cases, and the latter change does not occur in the absence of the first change, one change is the cause (or effect) of the other.

The effectiveness of the method of teaching, the enquiry approach, will be measured by way of calculating the rate developed by the teaching method. The rate is defined as the increment of growth in test performance occurring per unit interval of time.

The Sample

The primary schools sampled were only urban schools in Surabaya, Malang, and Yogyakarta. The schools sampled were (a) Surabaya – PPSP, SD Komparasi Ketintang, SD Komparasi Widoderen, and Santa Maria (catholic); (b) Malang – PPSP and Mardi Wiyata (catholic); and Yogyakarta – PPSP. Three classes of each school – grades 4, 5, and 6 – were tested. The number of students tested was 648 – 268 enquiry and 380 traditional students.

Method of Sample Selection

Samples were selected by the purposive or quota sampling technique. Quota sampling has been utilized because it is cheaper for a given study. In spite of the fact that quota sampling yields less precise results, it was felt that these would be satisfactory for the purpose desired.

The Delimitation of Variables

There are many variables in the sampled schools. They were divided into controlled and experimental variables. The controlled variables are ages of students, sex balance, socio-economic levels, parent education, parent occupation and medium of instruction. And the experimental variable is method of teaching.

The two groups are equated by randomization. This technique avoids the introduction of a bias into either the experimental or the control group.

The Instrument

This study used a norm-referenced test. It was a written, multiple-choice test with four alternatives – one correct answer and three distractors. The number of test items was 30.

Statistical Procedures

The rate of science concept development was measured statistically, using the method of least squares. It was based on three grades, 4, 5, and 6. The equation of a straight line used was $Y = a x + b$ ('a' is the rate).

Piagetian theory of cognitive development stated that cognitive structures (schemata) develop in a variant sequence. Many studies on Piagetian theory have indicated that cognitive schemata develop in a sigmoid curve. Since the curve will be used as a tool for an *ex post facto* analysis, the true rate was calculated by using the formula –

$$Wa = Wo\,e^{rt} \text{ ('r' is the rate)}$$

Limitations of this Study

This study is limited to (a) primary school; (b) the enquiry method as applied in PPSP schools; (c) specific area of science concepts; (d) the type of enquiry approach applied is the expository or guided discovery; and (3) the level of significance used is 0.05.

Results

Frequency distributions in Table 1 contain measures, by grade, of the number of students (N), mean scores (X), standard deviation (S), and standard error of the mean (S_x). Maximum score is 30. The numbers of students by grade, are even numbers, except Catholic Santa Maria; it is a girls' school. The even numbers are obtained by sex-balancing the students. This is done since sex differences do exist in some areas of studies (Miller, 1972, p. 5).

TABLE 1

Mean scores and standard deviations, by grade, teaching method, and medium of instruction of a sample of 848, December 1977.

School	*Grade 4*				*Grade 5*				*Grade 6*			
	N	*X*	*S*	S_x	*N*	*X*	*S*	S_x	*N*	*X*	*S*	S_x
I. PPSP Group												
1. Yogyakarta	20	18.8	3.4	0.8	28	19.1	3.9	0.7	28	22.0	3.6	0.7
2. Surabaya	34	20.0	3.3	0.6	32	23.5	3.2	0.6	52	21.6	3.8	0.5
3. Malang	24	20.5	2.8	0.6	32	21.5	3.5	0.6	18	21.2	3.0	0.7
Total	78	19.8	3.2	0.4	92	21.5	4.2	0.4	98	21.6	3.6	0.4
II. Rote Groups												
Govern. Schools												
SD Ketinteng	30	15.5	4.2	0.8	30	17.9	3.6	0.7	36	19.0	4.4	0.7
SD Widodaren	32	13.2	4.8	0.8	42	18.4	5.1	0.8	36	20.0	3.2	0.5
Total	62	14.3	4.6	0.6	72	18.2	4.6	0.5	72	19.5	3.8	0.4
Cath. Schools												
SDK Malang	24	20.5	2.6	0.6	22	21.5	3.5	0.7	18	21.2	3.0	0.7
SDK-St.Maria*	33	19.6	2.8	0.5	41	21.0	4.3	0.7	36	22.7	3.2	0.5
Total	57	20.0	3.0	0.4	63	21.2	4.0	0.5	54	22.2	3.3	0.4
Norm	197	17.3	4.8	0.3	227	19.8	4.9	0.3	224	21.1	4.1	0.2

*girl-students school.

In general, the higher the grades, the higher is the mean scores except in three schools, i.e.: PPSP Malang, and PPSP Surabaya. The mean scores, by grade, of the PPSP schools (enquiry group) are statistically higher than the Government schools (rote-learning group), but the difference is not statistically significant from those of Catholic schools (Table 2 and 3).

TABLE 2

Learning outcomes of the PPSP and rote-learning groups. December 1977.
(From Table 1)

Grade	*Enquiry*			*Rote-learning*			*Z*	*Significancy*
	N_1	X_1	S_1	N_2	X_2	S_2		
4	78	19.8	3.2	62	14.3	4.8	7.87	Significant
5	92	21.5	4.2	72	18.2	4.5	4.78	Significant
6	96	21.6	3.6	72	19.5	3.8	3.88	Significant
Total	266			206				

TABLE 3

Learning outcomes of the PPSP and Catholic schools
(From Table 1)

Grade	*PPSP*			*Catholic*			*Z*	*Significancy*
	N_1	X_1	S_1	N_2	X_2	S_2		
4	78	19.8	3.2	57	20.0	3.0	-0.37	not Significant
5	92	21.5	4.2	63	21.2	4.0	0.48	not Significant
6	96	21.6	3.6	54	22.2	3.3	1.06	not Significant
Total	266			174				

TABLE 4

Percent rate* of concept development
(From Tables 2 and 3)

School	*Grade*		
	4	*5*	*6*
PPSP	-	8.2	0.5
Government	-	24.0	6.9
Catholic	-	5.8	4.6

*Using formula Wa = wo e^{rt}

The difference between the two means (Z), by grade, are shown in Column Z. The level of significance used is 0.05.

The rate (in percent) of science concept development of PPSP schools is slower than of the rote-learning schools. In general, the percent rate of concept development rises sharply through grade 5, then levels off afterwards. In particular, learning outcomes of the PPSP schools are significantly higher than of the Government schools, but the percent rate of concept development of the former is slower than that of the latter. This phenomenon conforms to the Piagetian sigmoid curve of cognitive development. The growth does not take place at a constant rate but varies depending upon socio-cultural environment.

Linear regression of learning outcomes of the PPSP, Government, and Catholic schools are shown in Figure 1. They are based on the achievement measurement in grades 4, 5, and 6 only. The mean scores corresponding to grades below four and to those above six have been extrapolated rather than measured. Extrapolating these measurements is necessary because the younger students do not have the skills needed to take the written test.

All the regression lines referred to above are then combined into one, the regression line of the norm group, and in turn into a curve, representing the science concept growth in Indonesian students (Figure 2). This figure represents the cognitive growth of Indonesian children (urban students). The curve reaches the senescence phase at grade 5, while that curve in American students tends to level off at grades 8 and 9 (Gronlund, 1976, p. 394).

Analysis and conclusion

There seems to be an anomaly that the outcomes of the PPSP schools are much higher than the Government school while the percent rates are reversed. This anomaly can be explained by reference to Piagetian theory of cognitive development which states that cognitive structures develop in an invariant sequence. That is, the course is the same for all children although the ages of attainment may vary with intelligence and environment.

Thus, even though the PPSP students develop science concepts at a slower rate than do the Government school children, yet the enquiry approach is more effective than the rote method. The reasons are that (1) the trend – from grades four to six – is too short to be able to portray the whole S-shaped, or sigmoid, curve; and (2) assuming that Piaget's theory is valid, the sigmoid curve of the PPSP must have grown earlier through the grand phase than of the traditional. And the rate of grand phase of the PPSP must have been higher than of the Government schools since the senescence point of the PPSP – at grade 5 – is higher than that of the traditional schools.

Outcomes of Catholic school (using rote-learning method) are not statistically different from those of the PPSP but they are much higher as compared with the other rote-learning group, that is the Government school. Are the outcomes of the Catholic school caused by the teaching method? The answer is "no". Since socioeconomic levels, parent education, parent occupation, and the school setting (urban or rural) affect the levels of learning outcomes, it means that the teaching method is not the cause of learning process. They are not causal relationships, but 'sufficient conditions'. Insofar as a relation between two variables is not perfect, the relation is not causal (Hirchi & Selvin, 1971, p. 434). And a factor is called a 'sufficient condition' for learning if the presence is invariably followed by process (*Ibid*, p. 437). If an independent variable and a dependent variable are perfectly associated, no other independent variable is needed: that is, perfect association implies single causation, and less-

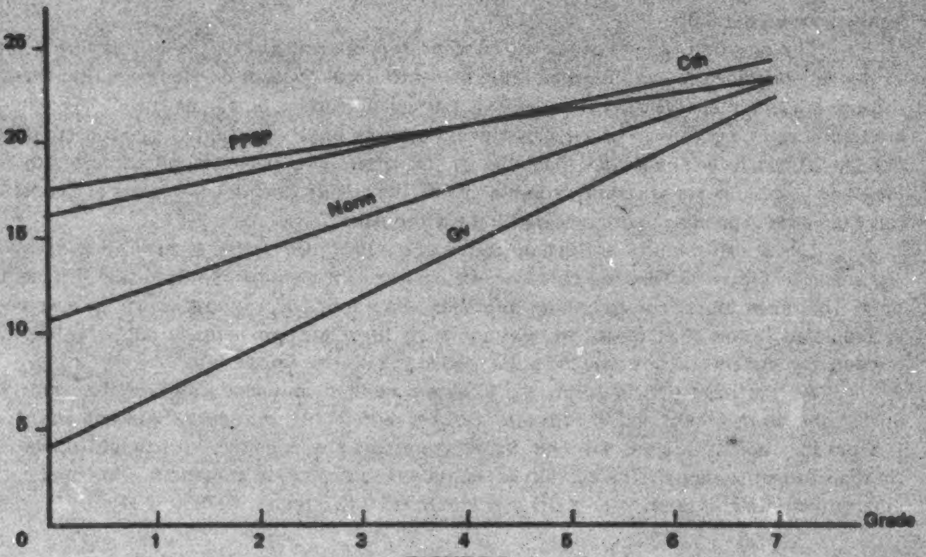


FIGURE 1

Regression lines of learning outcomes of the PPSP, Government (Gv), Catholic (Cth), and norm groups.

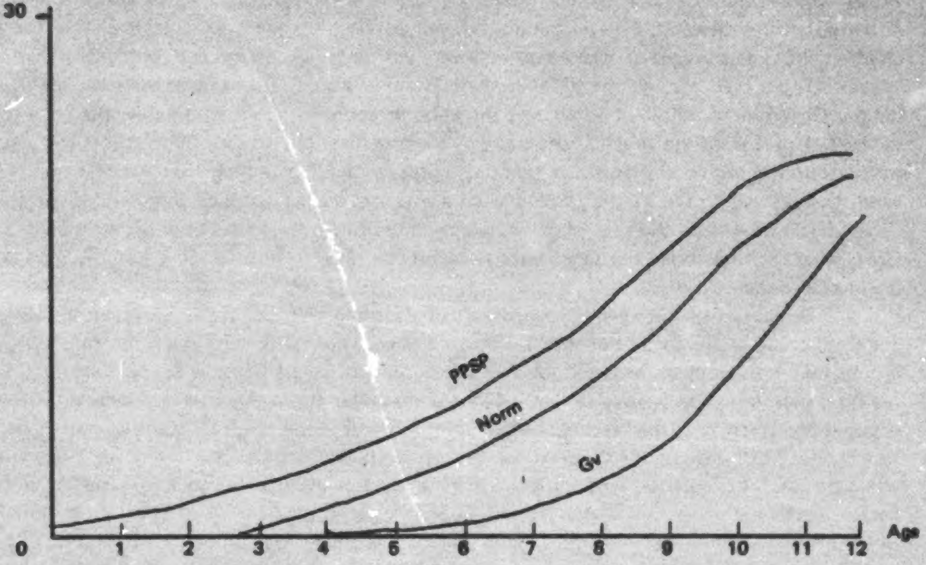


FIGURE 2

Science concept development in Indonesian (urban) children.

than-perfect association implies multiple causation.

The author concludes that the concomitant variables in the Catholic (urban) schools positively affect the learning process. The dominant influence comes from socioeconomic levels.

On the other hand, if the lower socioeconomic students learn by the enquiry method, the outcomes will be the same levels as the Catholic and are much higher as compared with those of the Government, rote-learning, school.

Summary of conclusions

1. The Indonesian student who learns science by the enquiry approach will develop science concepts at a higher rate than the student who learns by the rote method.
2. The western concept of teaching, the enquiry approach, is proved to be appropriate for younger children, below the age of nine, provided the enquiry method applied is the expository or guided discovery method.
3. In terms of the age levels of the achievement of formal operation, the delay in achieving science concepts in Indonesian children appears to be from 3 to 4 years, as compared with the normal western children.

References

COHEN, D. Evaluation of integrated science curricula. In P.E. Richmond (ed.), *New trends in integrated science teaching*. Paris: Unesco, 1973.

CRELLIN, C.T. Curriculum development in Indonesia, *Education in Asia*, 1976, 10, 1.

DeCECCO, J.P. & CRAWFORD, W.G. *The psychology of learning and instruction.* New Jersey: Prentice Hall, 1974.

GOOD, W.J. & HATT P.K. *Methods in social research.* New York: McGraw Hill, 1952.

GRONLUND, N.E. *Measurement and evaluation in teaching.* New York: Macmillan, 1976.

HIRCHI, T. & SELVIN, H.C. In Billy J. Franklin and H.W. Osborne (eds), *Research methods: Issues and insight.* Belmont: Wadsworth, 1971.

KAPLAN, A. In Billy J. Franklin and H.W. Osborne (eds), *Research methods: Issues and insight.* Belmont: Wadsworth, 1971.

LEWIS, L.J. Aims of science teaching in Africa. In P.G.S. Gilbert & M.N. Lovegrove (eds), *Science education in Africa.* London: Heinemann, 1972.

MAULDIN, L.W. *The colonial influences of Indonesian education.* Unpublished Ph.D. thesis – George Peabody College for Teachers.

MILLER, D.M. *Interpreting test scores.* New York: John Wiley and Son, 1972.

PRINCE, J.R. *Science concepts in a Pacific culture.* Sydney: Angus and Robertson, 1969.

RADCLIFF, D. Ki Hadjar Dewantara and the Taman Siswa School: Notes on an extra-colonial theory of education, *Comparative Education Review*, 1971, 15 (2), 219-226.

THOMAS, R.M. Course for writers of instructional modules. Unpublished paper: University of California, 1974.

Research in Science Education, 1978, 8, 48-57

LEVELS OF SCIENTIFIC ENQUIRY IN UNIVERSITY SCIENCE LABORATORY CLASSES: IMPLICATIONS FOR CURRICULUM DELIBERATIONS

Elizabeth H. Hegarty

Introduction

To those involved and interested in the processes of curriculum deliberation, it is encouraging to see that, alongside continuing widespread emphasis on theory development in science education, publications are beginning to build up in the realm which Joseph Schwab calls the "practical" (e.g. Reid & Walker, 1975; MacDonald & Walker, 1976). In contrast with the theoretic, the realm of the practical is concerned with choice and action in curriculum work and one of its main arts is deliberation.

Accounts of the practices of curriculum deliberations are now being published and include reports from course teams at the Open University (Ferguson, 1975; Harris, 1976; Lewis, 1971a,b,c; Newey, 1975; Riley, 1976) and from the Kettering Project at Stanford University (Eisner, 1975; Walker, 1971, 1975). These invite comparison with standards set in Schwab's position papers (1969, 1971, 1973, 1974) and permit recognition of a number of problem areas. Three of the problem areas detected in the published accounts are (1) difficulties with group behaviour in deliberation groups, (2) domination of the reported deliberations by subject matter concerns at the expense of the other curriculum "commonplaces" and (3) attention to the processes of curriculum planning at the expense of processes of curriculum implementation. Partial solutions offered for the first two problems include use of methods for structuring group processes (Hegarty, 1977, 1978), provision of special training programmes for curriculum teams (Sabar, 1977) and constitution of groups on the basis of Schwab's prescriptions for representation of the five commonplaces of curriculum deliberation: subject matter, learners, teachers, the milieu and the curriculum process (Fox, 1972; Hegarty, 1977, 1978; Sabar, 1977; Tamir and Jungwirth, 1974).

The processes of curriculum implementation (Johnson, 1967, 1969) were barely mentioned in some deliberation groups, whilst most attention was concentrated on the processes of curriculum planning (Lewis, 1971a, Walker, 1971, 1975). There is some suggestion that the same imbalance obtains in both deliberations on new curricula and deliberations on established curricula (Lewis, 1971c; Melton, 1977). Specifically, it would appear that groups have not used techniques for the systematic examination of prepared curriculum materials or for the observation of students using the materials, although of course research studies by individuals have employed such techniques. Eisner (1975) called for the situation to be remedied by future deliberation groups.

This paper presents part of a study which focussed on the work of groups responsible for the planning and revision of the laboratory work component of science courses at university level (especially microbiology). Reference will be made to content analysis of laboratory manuals, to observation studies of behaviour of staff and students in laboratory classrooms, and to description of the relationship between level of scientific enquiry of laboratory manuals and classroom behaviours.

Content analysis of science laboratory manuals

A description of content analysis which aptly summarizes the approach taken in the present study was provided by Clarke (1973) who saw it as a technique for evaluating the potential of written materials to allow students to achieve certain objectives. In a more general sense, content analysis has been described by Holsti (1969) in his comprehensive book as a valuable unobtrusive technique for investigating problems in which the content of communication serves as a basis for inference.

Of the few studies broadly relevant to science curricula, some are directed towards making inferences as to the effects of communication (for example Gardner's study, 1968, of the readability of the PSSC text). Most others basically describe characteristics of communications and proceed to check content against implicit or explicit standards. These standards might involve comparisons between the stated goals of writers and the materials they produce, e.g. the study of levels of scientific enquiry in laboratory manuals for PSSC, BSCS and CHEM study (Herron, 1971). Other standards involve the appropriateness of certain styles and structures for instructional materials, e.g. analysis of the "style" of ASEP (Shepherd, 1972) and the pre-trial formative evaluation of the content and structure of an ASEP unit (Clarke, 1973). Still other standards involve judgement of the appropriateness of demands for certain intellectual skills and reasoning abilities (Airasian, 1970; Haley and Good, 1976; Ingle and Shayer, 1971; Shayer, 1970, 1974).

In the present study, two schemes were used for content analysis of laboratory manuals used in microbiology in university science. One concerning the *level* of scientific enquiry was developed by Herron (1971) based on the theoretical writings of Schwab (1962, 1964). The second, concerning the *types* of enquiry behaviours demanded was specially developed for the study, based on the work of Schwab (1964), Herron (1971) and Klopfer (1971). The Item (= entire laboratory exercise) was chosen as the correct unit of analysis (Holsti, 1969; Kerlinger, 1973) and explicitly formulated rules were developed for selection of Items and for coding procedures. Results were obtained after pooling the analyses of a team of four coders, between pairs of whom Scotts pi coefficients of inter-coder agreement ranged from 0.72 to 0.92 (Holsti, 1969; Scott, 1955).

Herron's scheme for determination of levels of scientific enquiry is shown in Table 1 (modified at level 2). It is a type of Guttman scale with increasing degrees of openness for students to choose the aims and methods of their laboratory work. At the lowest level one might find exercises which were simply practice in techniques or confirmatory exercises where the answer was already provided for students. At the highest level one finds laboratory work where students have freedom to determine the nature of the problem on which they work as well as designing materials and methods they use.

It should be pointed out that the notion of levels of scientific enquiry is value laden, having the potential for what might be called ideological commitment – where one assumes that higher levels are better than lower levels. Nonetheless, it would seem reasonable that university students in final year of their major subject should experience scientific enquiry at the highest levels.

One would expect to support this view by using arguments grounded in the philosophy of science. But encouraging empirical support comes from research studies carried out at university level by Holcomb (1971) and Wolfson (1976). These authors found that as the degree of guidance in laboratory work decreased (and student design of procedures increased), so students' retention of what they had learned increased.

Level 0 Items present a problem. Practice in techniques can easily be justified, though not necessarily in a final year course (Newman, 1978, called such practice "pre-science", initiation into real science), but the use of "recipes" where the answer is already well known

TABLE 1

Level of scientific enquiry in microbiology laboratory manuals

Table shows definition of levels and percentages of laboratory exercises at different levels. Extended exercises and projects are excluded.

Level	*Aim*	*Materials*	*Method*	*Answer*	*UNSW (N=63)*	*SU (N=78)*
0	Given	Given	Given	Given	33.3	34.6
1	Given	Given	Given	Open	38.1	62.8
2A	Given	Given whole or part	Open or part given	Open	25.4	1.3
2B	Given	Open	Open	Open	0	0
3	Open	Open	Open	Open	0	0
Unknown	Evidence lacking				3.2	1.3

$\chi^2 = 23.2$, df = 2, $p < 0.001$

to staff and students, is difficult or impossible to justify. In terms of the analysis by Newman (1978), recipes are "copying" and are not science at all.

Table 1 also shows results for content analysis of levels of scientific enquiry in laboratory manuals for final year microbiology at two Australian universities (UNSW = University of New South Wales, SU = University of Sydney). The manuals are used in final year courses which constitute a major sequence in microbiology. Both have similar percentages of Items at level 0 (about 34%) but UNSW has 25% of Items at level 2 whereas SU has only 1%. This analysis was restricted to standard "Items", i.e. exercises which might be completed in one or two class periods (of 3 to 4 hours each). Extended exercises or projects were analyzed separately and results showed that one such project was used at SU (level 2B) and that there were six at UNSW chiefly at levels 2A and 2B.

Comparison of the microbiology data for standard Items with Herron's data at school level showed an encouraging shift away from the use of curriculum materials at level 0 (about 75% in the school study) and, for microbiology at UNSW (but not SU) an encouraging use of laboratory curriculum materials at level 2, congruent with the central role attributed to processes of scientific enquiry by many writers.

The specially developed content analysis scheme for determining the types (rather than levels) of scientific enquiry behaviours is shown in abbreviated form in Table 2, together with the results of analysis for microbiology laboratory manuals at UNSW and SU. For each Item, the maximum score per category is 1, but each Item may be scored in more than one category. Cross checking the two schemes showed internal consistency, in that no Items classified as level 0 received scores for enquiry behaviours. A score for category 7 (Analysis and Interpretation) was the most common indicator for items at level 1 or above.

Excluding the high scores of category 7, only 4% of Items at SU and 18% at UNSW demanded more than one type of enquiry behaviour in each laboratory exercise, and as can be seen from the Table, three categories received no scores at all (Nos. 1, 2, 10). Taking into account the sketchy evidence provided about projects, it appears unlikely that these three

TABLE 2

Types of scientific enquiry in microbiology laboratory manuals

Excludes extended exercises and projects

Category	UNSW (N=63)	SU (N=78)
1. Identifies problem	0	0
2. Chooses from problems	0	0
3. Formulates hypothesis	7.2%	0
4. Designs investigation	12.3%	3.5%
5. Designs technique	5.6%	1.8%
6. Selects technique	25.0%	6.9%
7. Analyses. Interprets	44.2%	55.2%
8. Evaluates Hypothesis	2%	0
9. Discusses implications	14.1%	9.3%
10. Revises conceptualisations	0	0

enquiry behaviours were demanded in any part of the laboratory course at either institution.

It may be concluded that the overall higher levels and emphasis on scientific enquiry at UNSW was reflected in the positive, albeit low scores for formulation and evaluation of hypotheses, two of the central processes of scientific enquiry. In addition, there were reasonably high scores for the design of experimental investigations to test hypotheses and in the design or selection of necessary laboratory methods or techniques.

Observations of behaviour in science laboratory classrooms

In the research tradition of classroom interaction, relatively few of the studies published to date involve science classrooms, even fewer involve science laboratory classes and science classrooms at universities and colleges have not been studied at all – despite the fact that they often represent the only opportunity for staff-student interaction. The rather scanty literature resulting from the observation of science classrooms has been reviewed by Tisher (1972) and very recently by Power (1977) in the context of developments in the field of classroom interaction studies as a whole.

Most of the research was influenced by a desire to develop a coherent theory of teaching. Disappointingly little success was obtained but Power's review (1977) points up ways in which research in the realm of the theoretic may become productive in the future. Moves away from the realm of the theoretic to the realm of the practical are beginning. Arguments for the use of interaction analysis in curriculum development and curriculum evaluation have been presented recently by Eggleston and coworkers (1975, 1976). The conceptual approach of Good and Power (1976) to designing successful classroom environments for different types of students will inform the thinking of workers in both realms, theoretic and practical.

The first part of the classroom research to be presented involved a baseline study to determine the nature of interactions in some university science laboratory classrooms since these have been little studied.

Since the behaviours of students and staff are likely to be the subject of separate consideration at curriculum deliberation meetings, seperate observations of students and tutors were made, using two slightly different versions of the specially developed System for Observation of Laboratory Behaviours of Students/Tutors (abbreviated to SOLS and SOLT respectively). A new system was needed because of defects in previous systems used for observation of laboratory classes (especially lack of mutual exclusiveness of categories), although availability of the recently published SABIC scheme (Power and Tisher, 1976) may have avoided the need for development of a new scheme. Live observation was used rather then videotaping because of problems of noise and the mobility of students and tutors when all are present simultaneously in one large university laboratory. Under such circumstances, Dunkin and Biddle (1974, Ch. 4) suggest that two facets are probably the maximum manageable and the two used in the present study were Behaviour and Target. Ten second time intervals were used as the unit of analysis. Data were collected by three different trained observers and coefficients of inter-observer agreement were 0.8 or above, using Scott's pi (Mitchell, 1969; Scott, 1955).

Observations were made during 18 classes (approximately three quarters of all the laboratory classes) scheduled during a one semester course in General Microbiology for microbiology majors in their final year of studies at UNSW and also during the corresponding course at SU. Classes were of four hours duration at UNSW and three hours at SU.

Table 3 shows the results with categories grouped into five gross divisions.

TABLE 3

Behaviours in microbiology laboratory classes

Figures in the table represent percentages of total behaviours recorded

Division	*Student Study*[1]		*Tutor Study*[2]	
	UNSW	*SU*	*UNSW*	*SU*
I Cognitive	26.6	30.3	41.2	61.1
II Laboratory activities	28.5	30.1	14.3	4.7
III Laboratory organization	28.5	28.4	29.3	28.1
IV Class unrelated	15.6	10.6	13.2	5.9
V Other	0.8	0.6	2.0	0.2
Total	100%	100%	100%	100%
N	14670	8520	21500	6670

1. $\chi^2 = 134$ (df = 4) $p < 0.001$
2. $\chi^2 = 1150$ (df = 4) $p < 0.001$

The findings are not intended as a Method A vs. Method B comparison but are presented to give an account of the commonplaces of laboratory behaviour in microbiology classes at university level. It was found that at both institutions students' time was divided approximately equally between the three main divisions, I Cognitive, II Laboratory activities, III Laboratory organization. The main difference between institutions occurred on behaviour unrelated to classwork, of which 15.6% of time was spent at UNSW and 10.6% at SU. When tutors' behaviour was observed there were fewer obvious similarities between institutions although Division III, Laboratory organization, was constant at 28-29%. Otherwise there was, as with students, far more behaviour apparently unrelated to classwork at UNSW and there were marked differences in scores for Divisions I and II. The microbiology tutors studied, especially those at SU, apparently spent far more time on cognitive behaviours (40-60%) than teachers in any school study (where the highest was 33%, Evans and Balzer, 1970). This difference was mostly attributable to heavy emphasis on talk about laboratory procedures, and (at SU) a great deal of talk at low cognitive levels.

When the relative proportions of verbal and non-verbal behaviours of students and tutors were examined it was found that the results for students at both institutions and for tutors at UNSW are similar to those reported at school level (Power, 1977). Findings for tutors at SU were quite different from any previously reported in science laboratory classes. Tutors at SU spent the majority of their time talking (72%) whereas figures in the range 35-60% are the highest reported elsewhere. This does not mean that tutors dominate the activities of the whole classroom, except during addresses with the assembled group or class as target (12%) but it does mean that "tutor talk" best describes the nature of transactions with individual students or small groups.

Effect of levels of enquiry of microbiology laboratory manuals on the behaviour of students and tutors in laboratory classes

Comparisons have been made of the behaviour of teachers at school level using different curriculum materials, e.g. BSCS (enquiry oriented curriculum) vs. "non BSCS" (Balzer, 1970). However from the review by Power (1977) of factors influencing the nature of science classroom interactions, it appears that no studies have been made of laboratory behaviours before and after the introduction of enquiry oriented curriculum materials nor of the effect of curriculum materials at different levels of enquiry on behaviour within a single course at either school or university level.

The present sutdy was designed to investigate the effects of different levels of scientific enquiry of microbiology laboratory manuals on the behaviour of students and tutors in laboratory classes. The scheme developed by Herron (discussed earlier) was used to determine the level of scientific enquiry of laboratory manuals and the study was restricted to UNSW since the range of levels at SU was narrow.

As noted earlier, no laboratory exercises at level 3 were reported in the school or university studies. Thus, in practice, Herron's scale has 3 levels, 0, 1, 2. For the purposes of the present study, these were regarded as Low, Medium and High levels of enquiry and levels 0 and 2 were selected for further investigation. The expectation was that as the level of enquiry of manuals increases from Low to High, so certain differences in laboratory behaviours of students and staff would be observed (99.9% confidence limits were used to check differences). The full results of changes with level of enquiry in all 20 odd categories in Facet 1 (Behaviour) of the two observation schemes are shown in Tables 4 and 5. A summary follows of some of the more interesting results.

As the level of enquiry increased, there was a marked increase in the proportion of tutors' time spent on the development of substantive content, especially at low cognitive levels, and a concomitant marked increase in students' time spent listening to such talk. There was a slight trend towards a similar increase in the proportion of students' time spent on talk at low cognitive levels. More importantly, there was a significant increase in time spent by students talking on scientific processes, most of which was accounted for by talk with a fellow student not a tutor, as target. Results for tutors were quite different, however. Time spent on

TABLE 4

Effect of level of enquiry of laboratory manuals on student behaviours in microbiology laboratory classes at UNSW

Figures in the table represent percentages of total student behaviours recorded at each level.

Division and categories of SDLS		*Level of enquiry*	
	0	*2A*	*2B*
I. *Cognitive*			
Content			
1.1 Listen	0.5	3.8	1.2
1.2 Listen (interaction)	0.7	4.7	2.8
1.3 Talk – low cognitive	5.6	9.2	6.5
1.4 Talk – scientific processes	0.1	3.0	1.0
Content total	6.9	20.7	11.5
Procedures			
1.5 Listen	9.9	4.7	1.7
1.6 Listen (interaction)	2.4	1.4	4.1
1.7 Talk – procedures	6.6	8.9	6.4
Procedures total	18.9	15.0	12.2
Total	25.8	35.7	23.7
II. *Laboratory activities*			
2.1 Techniques	27.2	18.7	16.7
2.2 Observations	14.7	4.2	5.3
Total	41.9	22.9	22.0
III. *Laboratory organization*			
3.1 Talk	1.1	1.1	0.8
3.2 Organization	11.6	11.4	21.4
3.3 Reads	5.0	9.2	6.7
3.4 Writes	5.4	4.7	4.2
Total	23.1	26.3	33.0
IV. *Unrelated*			
4.1 Unrelated interaction	0.9	3.7	3.5
4.2 Missing	3.3	6.7	9.5
4.3 Unrelated activity	3.9	4.0	8.0
Total	8.1	14.4	20.9
V. *Other*			
5.1 Affective: listen	0.6	0.3	0.2
5.2 Affective: talk	0.0	0.0	0.0
5.3 Unclassifiable. Missing data	0.6	0.3	0.3
Total	1.2	0.6	0.5
Total	100	100	100
Observations at each level, N =	4410	3420	6480

TABLE 5

Effect of level of enquiry of laboratory manuals on tutor behaviours in microbiology laboratory classes at UNSW

Figures in the table represent percentages of total tutor behaviours recorded at each level.

Divisions and categories of SOLT		*Level of enquiry*	
	0	*2A*	*2B*
I. *Cognitive*			
Content			
1.1 Talk – low cognitive	1.5	7.9	7.0
1.2 Talk – scientific processes	4.1	5.1	1.0
1.3 Listen	2.2	5.0	5.8
Content total	7.8	18.0	13.9
Procedures			
1.4 Talk – procedures	24.7	23.0	12.8
1.5 Listen	8.2	6.5	7.2
Procedures total	32.9	29.5	20.0
Total	(40.7)	(47.5)	(33.9)
II. *Laboratory activities*			
2.1 Assist with equipment, resources	1.7	2.7	2.3
2.2 Give demonstrations	4.4	0.8	2.5
2.3 Watch student working	4.3	3.8	6.3
2.4 Examine student materials	7.4	3.5	5.1
Total	17.8	10.8	16.2
III. *Laboratory organization*			
3.1 Talk – locating materials	2.3	2.4	1.6
3.2 Talk – laboratory management	5.8	6.5	3.3
3.3 Talk – teaching methods	0.7	0.9	0.3
3.4 Listen	5.5	3.0	3.7
3.5 Supervision	2.2	2.2	3.8
3.6 Laboratory management	10.6	11.1	10.6
3.7 As 3.6, not at scheduled location	1.3	3.0	8.5
Total	28.4	29.1	31.8
IV. *Unrelated*			
4.1 Unrelated interaction	5.5	4.0	4.5
4.2 Missing	2.8	5.3	4.2
4.3 Unrelated activity	3.9	1.6	7.1
Total	12.2	10.9	15.8
V. *Other*			
5.1 Unclassifiable	0.6	0.3	1.2
5.2 Data missing	0.3	1.6	0.8
Total	0.9	1.9	2.0
Total	100%	100%	100%
Observations at each level, N =	4860	8640	5310

scientific processes did increase at level 2A, but during extended exercises and projects at level 2B, the talk on scientific processes dropped below the score at level 0.

One of the most noticeable effects on tutors' behaviour was an increase in laboratory management responsibilities - not management within the laboratory, which remained constant - but in management activities not at any scheduled location, i.e. we have the staff away from the students altogether, missing from the classrooms, running around the building to preparation areas and private research laboratories collecting and organizing materials for students' projects. During level 2B projects, the total proportion of tutors' time spent on management activities was 20%. Similarly, there was a marked increase in time spent by students on laboratory organization, about 20% at level 2B but not at lower levels. It would appear that as students were required to spend more time planning and organizing laboratory work at higher levels of scientific enquiry, less time was taken for using laboratory techniques, making observations and taking measurements. One final effect was unexpected, though perhaps should not have been since Power (1977) recently reported similar findings at school level. This was the increase in the proportion of students' time spent on talk and other activities apparently unrelated to classwork. This effect was particularly marked during laboratory work at level 2B.

Discussion

This paper has presented results of a study directed towards improving the repertoire of information available to curriculum deliberation groups. In a sense then its value should be judged in terms of its relevance to the curriculum groups responsible for the microbiology courses investigated. Before one could be confident of generalizing further, what is needed is a large body of reports on the information used by curriculum groups, and on the judgements resulting from the deliberations of the groups. Only then will we have the grounds for testing our views on what constitutes better and worse in the art of deliberation on science curricula (Schwab, 1978).

A report has been compiled of the deliberations conducted in the School of Microbiology at UNSW but it is difficult to summarize it within the limits of this paper. The discussion will therefore be limited to general issues illuminated as far as possible by judgements from the microbiology curriculum deliberation group.

The arts of the practical are likely to help workers arrive at better choices and, to quote Schwab again "in a probabilistic world, we must be satisfied with that" (Schwab, 1974). So curriculum workers should be prepared to make tactical judgements involving use of data based on frequency counts in the absence of absolute standards. These judgements would be tempered by the warnings of McClellan (1971) who pointed out that cognitive rationality is a matter of form not frequency, and of Power (1977) who explained that it would be foolish to expect direct relationships between frequency counts and outcomes.

When enquiry oriented curriculum materials have been designed, a first step is to check that the materials reflect the designers' aims. Thus, content analysis of BSCS and PSSC manuals were judged by Herron (1971) and Westbury (1970) to indicate that scientific enquiry was preached by the curriculum designers a deal more often than it was practised. By comparison, the curriculum group at UNSW felt that the shift towards upper level exercises was encouraging but that action should be taken to reduce the figure of 30% of set exercises at level 0 still lower. The second content analysis scheme provided guidance as to which enquiry behaviours should be emphasised when replacement exercises were being developed.

An aim of introducing enquiry-oriented curricula is to encourage an increase in behaviours congruent with scientific enquiry (Eggleston & Galton, 1976) and it is often assumed desirable that such an increase should be manifested in the enquiry behaviour of both students and staff (Balzer, 1969; Eggleston & Galton, 1976).

Because of the original emphasis on scientific enquiry, some findings did not seem immediately controversial - but were very interesting in their implications - the increase in time spent by tutors on laboratory management activities outside the classroom and away from students was an unforeseen and presumably undesirable result of curricular change. This finding would seem worthy of debate by curriculum deliberators.

Other findings were more difficult to interpret. What are the implications of the finding that with increase in level of scientific enquiry, the increase in time spent by students talking on scientific processes was not matched by an increase in time spent by tutors?

Evidence from interviews suggested that university tutors often sought a role for themselves as resource persons; especially so during classes at higher levels of enquiry, rather than acting as a challenger or promoter of enquiry behaviours. This attitude is perhaps due to incorrectly equating scientific enquiry with (unguided) student discovery (an elision which has been fully discussed by Keislar and Shulman, 1966). There is an alternative interpretation in accord with an organization model developed at the Tavistock Institute of Human Relations (Miller, 1976). When the overall purpose of the course changes to emphasize scientific enquiry, the role of tutors is changed correspondingly. If tutors do not possess the relevant new skills, they will not be able to discharge the authority of their roles. They and their roles will therefore seem diminished with respect to their students. Under such circumstances of perceived threat, increase of concentration on management activities (especially away from the classroom) might be a predictable response. Deliberators might debate the acceptability of tutors behaviour such as that reported in the present study and elsewhere (Power and Tisher, 1976). Opportunities could be canvassed for staff development programmes to promote an understanding of the nature of scientific enquiry and the acquisition of relevant teaching skills.

On the other hand, it is also often assumed desirable that the introduction of enquiry-oriented curricula should be accompanied by a decrease in students' dependence on the teacher and an increase in peer group interaction (Egelston, 1973; Nelson, 1973). Opportunity for students to compare observations and discuss interpretations is thought to facilitate self-regulation (in Piagetian terminology) and the acquisition of abstractions (Piaget, 1964; Nelson, 1973; Lawson & Renner, 1975). It was encouraging then that in the present study, increase in the level of enquiry of microbiology laboratory manuals was associated with an increase in the proportion of students' time spent in discussion with other student/s as target.

Finally, it is worth mentioning that enquiry related behaviours of both students and staff were more noticeable at level 2A than at level 2B, and that at level 2B the most prominent behaviours involved organization and management. These findings represent a challenge to the practice of offering project work in university science. In view of the notorious expense of time and money on projects, deliberators might debate the replacement of some projects with structured exercises at level 2A (which apparently present more opportunities for enquiry within a compressed time span).

With this discussion of the value of content analysis and classroom observation data to curriculum deliberators, has come a sense of the priorities which must be weighed, the compromises and tradeoffs which are involved in replanning curricula. But to reiterate, in order to improve our use of the arts of the practical, what is needed is a large body of such reports on the information used by curriculum groups and on the judgements resulting from their deliberations.

References

AIRASIAN, P.W. The use of heirarchies in the analysis and planning of chemistry instruction. *Science Education*, 1970, 54, 91-95.

BALZER, L. Non-verbal and verbal behaviours of biology teachers. *American Biology Teacher*, 1969, 31, 226-229.

BALZER, L. Teacher behaviors and student enquiry in biology. *American Biology Teacher*, 1970, 32, 26-28.

CLARKE, J.A. The role of the content and structure of curriculum material in cognition. *Science Education: Research 1973*, 1973, 119-141.

DUNKIN, M.J. & BIDDLE, B.J. *The study of teaching.* New York: Rinehart and Winston, 1974.

EGELSTON, J. Inductive vs. traditional methods of teaching high school biology laboratory experiments. *Science Education*, 1973, 57, 467-477.

EGGLESTON, J.F., GALTON, M.J. & JONES, M.E. *A science teaching observation schedule.* London: Macmillan Education, 1975.

EGGLESTON, J. & GALTON, M. Curriculum evaluation and interaction analysis. *British Journal of Teacher Education*, 1976, 2, 189-198.

EISNER, E.W. Curriculum development in Stanford University's Kettering project: recollections and ruminations. *Journal of Curriculum Studies*, 1975, 7, 26-41.

EVANS, T.P. & BALZER, L. An inductive approach to the study of biology teacher behaviors. *Journal of Research in Science Teaching*, 1970, 7, 47-56.

FERGUSON, J. *The open university from within.* London: University of London Press, 1975.

FOX, S. A practical image of "the practical". *Curriculum Theory Network*, 1972, 10, 45-57.

GARDNER, P.L. The P.S.S.C. text: a measure of its readability. *Australian Science Teachers' Journal*, 1968, 14(3), 71-74.

GOOD, T.L. & POWER, C.N. Designing successful classroom environments for different types of students. *Journal of Curriculum Studies*, 1976, 8, 45-60.

HALEY, S.B. & GOOD, R.G. Concrete and formal operational thought: implications for introductory college biology. *The American Biology Teacher*, 1976, 38, 407-412, 430.

HARRIS, D. Educational technology at the Open University: a short history of achievement and cancellation. *British Journal of Educational Technology*, 1976, 7(1), 43-53.

HEGARTY, E. The problem identification phase of curriculum deliberation: use of the nominal group technique. *Journal of Curriculum Studies*, 1977, 9, 31-41.

HEGARTY, E. Problems associated with current full-time, degree-level educational programmes for medical laboratory personnel I. Procedures used to identify problems. *Australian Journal of Medical Technology*, 1978, 9, 7-12.

HERRON, M.D. The nature of scientific enquiry. *School Review*, 1971, 79, 171-212.

HOLCOMB, C.M. The effect of degrees of direction in the qualitative analysis laboratory on retention of learning. *Journal of Research in Science Teaching*, 1971, 8, 165-169.

HOLSTI, O.R. *Content analysis for the social sciences and humanities.* Reading, Massachusetts: Addison-Wesley, 1969.

INGLE, R.B. & SHAYER, M. Conceptual demands in Nuffield O-level chemistry. *Education in Chemistry*, 1971, 8, 182–183.

JOHNSON, M. Definitions and models in curriculum theory. *Educational Theory*, 1967, 17, 127–140.

JOHNSON, M. The translation of curriculum into instruction. *Journal of Curriculum Studies*, 1969, 1, 115–131.

KEISLAR, E.R. & SHULMAN, L.S. The problem of discovery: conference in retrospect. In L.S. Shulman & E.R. Keislar (eds), *Learning by discovery: a critical appraisal*. Chicago: Rand McNally, 1966.

KERLINGER, F.N. *Foundations of behavioral research*. (2nd ed.). New York: Holt, Rinehart and Winston, 1973.

KLOPFER, L.E. Evaluation of learning in science. In B.S. Bloom, J.T. Hastings & G.F. Madaus (eds), *Student learning*. New York: McGraw Hill, 1971.

LAWSON, A.E. & RENNER, J.W. Piagetian theory and biology teaching. *The American Biology Teacher*, 1975, 37, 336–343.

LEWIS, B.N. Course production at the Open University I: some basic problems. *British Journal of Educational Technology*, 1971a, 2(1), 4–13.

LEWIS, B.N. Course production at the Open University II: activities and activity networks. *British Journal of Educational Technology*, 1971b, 2(2), 111–123.

LEWIS, B.N. Course production at the Open University III: planning and scheduling. *British Journal of Educational Technology*, 1971c, 2(3), 189–204.

MacDONALD, B. & WALKER, R. (eds). *Changing the curriculum*. Melbourne: Macmillan, 1976.

McCLELLAN, J.E. Classroom-teaching research: a philosophical critique. In I. Westbury & A.A. Bellack (eds), *Research into classroom processes: recent developments and next steps*. New York: Teachers College Press, 1971.

MELTON, R.F. Course evaluation at the Open University: a case study. *British Journal of Educational Technology*, 1977, 8(2), 97–103.

MILLER, E.J. (ed.). *Task and organization*. London: Wiley, 1976.

MITCHELL, J.V. Education's challenge to psychology: the prediction of behaviour from person-environment interactions. *Review of Educational Research*, 1969, 39, 695–721.

NELSON, M.A. Discussion strategies and learning science principles. *Journal of Research in Science Teaching*, 1973, 10, 25–38.

NEWEY, C. On being a course team chairman. *Teaching at a Distance*, 1975, 1(4), 47–51.

NEWMAN, B.S. The scientific status of any so-called science. *Research in Science Education*, 1978, 8, 1–9.

PIAGET, J. Development and learning. *Journal of Research in Science Teaching*, 1964, 2, 176–186.

POWER, C. A critical review of science classroom interaction studies. *Studies in Science Education*, 1977, 4, 1–30.

POWER, C.N. & TISHER, R.P. Relationships between classroom behavior and instructional outcomes in an individualized science program. *Journal of Research in Science Teaching*, 1976, 13, 489–497.

REID, W.A. & WALKER, D.F. (eds). *Case studies in curriculum change*. London: Routledge and Kegan Paul, 1975.

RILEY, J. Course teams at the Open University. *Studies in Higher Education*, 1976, 1(1), 57–61.

SABAR, N. An integrative in-service training for a curriculum team. *Journal of Curriculum Studies*, 1977, 9, 81–89.

SCHWAB, J.J. The teaching of science as enquiry. In J.J. Schwab & P.F. Brandwein (eds), *The teaching of science.* Cambridge, Massachusetts: Harvard University Press, 1962.

SCHWAB, J.J. The structure of the natural sciences. In G.W. Ford & L. Pugno (eds), *The structure of knowledge and the curriculum.* Chicago: Rand McNally, 1964.

SCHWAB, J.J. The practical: a language for curriculum. *School Review*, 1969, 78, 1–23.

SCHWAB, J.J. The practical 2: arts of eclectic. *School Review*, 1971, 79, 493–542.

SCHWAB, J.J. The practical 3: translation into curriculum. *School Review*, 1973, 81, 501–522.

SCHWAB, J.J. Decision and choice: the coming duty of science teaching. *Journal of Research in Science Teaching*, 1974, 11, 309–317.

SCHWAB, J.J. Foreword in W.A. Reid & D.F. Walker (eds), *Case studies in curriculum change.* London: Routledge and Kegan Paul, 1975.

SCOTT, W.A. Reliability of content analysis: the case of nominal coding. *Public Opinion Quarterly*, 1955, 19, 321–325.

SHAYER, M. How to assess science courses. *Education in Chemistry*, 1970, 7, 182–184, 186.

SHAYER, M. Conceptual demands in the Nuffield O-level biology course. *School Science Review*, 1974, 56, 381–388.

SHEPHERD, S.R. A checklist for analysing the 'style' of instructional materials. *Research 1972*, 1972, 28–33.

TAMIR, P. & JUNGWIRTH, E. Botany and zoology – a curriculum problem. *Journal of Research in Science Teaching*, 1974, 11, 5–16.

TISHER, R.P. Classroom interaction analysis: the new religion. *Research 1972*, 1972, 35–49.

WALKER, D.F. Strategies of deliberation in three curriculum development projects. Unpublished Ph.D. thesis, Stanford University, 1971.

WALKER, D.F. Curriculum development in an art project. In W.A. Reid, & D.F. Walker (eds), *Case studies in curriculum change: Great Britain and the United States.* London: Routledge and Kegan Paul, 1975.

WESTBURY, I. Curriculum evaluation. *Review of Educational Research*, 1970, 40, 239–260.

WOLFSON, M.L. Lectures versus individual research: their effect on retention. *Journal of College Science Teaching*, 1976, 6(1), 36–37.

Research in Science Education, 1978, 8, [illegible]

THE PRODUCTION OF CRITERIA FOR EVALUATING SCIENCE CURRICULUM MATERIALS WITH SPECIFIC REFERENCE TO CRITERIA FOR INQUIRY AND THE APPLICATION OF THESE CRITERIA TO CHEMISTRY CURRICULUM MATERIALS

D.A. Mitchell

Introduction

Recent developments in Queensland have jolted curriculum developers and evaluators as well as teachers and the public into the need for careful scrutiny of new curriculum materials before they are adopted. Wolf (1969), in a tongue-in-cheek article, critizes the "cosmetic method" in curriculum evaluation. Essentially, this involves taking a cursory look at a program and deciding if it looks good. Wolf refers also to a "cardiac" approach which is used to resolve dilemmas presented by data obtained from some systematic empirical approach. All one must do here is dismiss the data and believe in one's heart that the new program is indeed a good one. These approaches are not good enough: nor are three others mentioned by Wolf adequate, viz. the "colloquial" method, the "curricular" method or the "computational" method.

Evaluation of curriculum materials themselves, based on systematic analysis of content and activities has rarely been undertaken (Bloom, 1966; Herron, 1971; Payne, 1974; Champagne and Klopfer, 1974). This happened in spite of warnings that evaluation must be seen as part of the development process, "the other side of the curriculum coin of curriculum development" (Bebell, 1962, p. 4).

In this paper an evaluation study based on an analysis of content and activities of a chemistry program is described. An investigation of the extent to which a secondary school chemistry curriculum package, purporting to emphasize inquiry, includes the concept of inquiry was undertaken. Inquiry was seen from two separate, but reconcilable, indeed, mutually supportive points of view, viz. inquiry in chemistry and inquiry as a pedagogical principle. A table of specifications for inquiry which was developed by Klopfer (1971) was used as a basis for a coding manual for analysis of the curriculum package *Chemistry, Experimental Foundations* (Parry, *et al.*, 1970) which is a recent version of the CHEM Study curriculum materials, the essential components of which are the text, the laboratory manual, and the teachers' guide.

Inquiry

Lucas (1971, p. 25) states that inquiry can be used to refer to

(a) what scientists do in obtaining answers from nature (their techniques and procedures)

(b) the logical processes used in science

(c) the teaching technique that
   (i) enables the pupil to solve problems by asking questions to gather information (Suchman)
   or

(ii) uses a semi-structured discussion intended to develop skill in probing or searching (Schwab).

(d) a combination of the meanings using the technique (meaning 'c') to teaching about inquiry (meaning 'a').

Inquiry in chemistry

Chemistry is a discipline with substantive and syntactical structure. It provides for its members a disciplined form of intellectual satisfaction of curiosity. In pursuit of inquiry a range of methods is available and is used. Chemical knowledge is subject to continuous development and updating. Inquiry and discovery in chemistry, even accidental discovery, requires a prepared mind. Much of the work of chemists is routine, painstaking, detailed investigation. Theory and experiment each has an important place and the two interact continuously. The molecular basis of chemistry provides many of the current paradigms for inquiry. Chemistry has a set of important unifying themes that subsume concepts and generalizations and allow for the organization of a vast array of facts. In experimental investigation, chemists rely heavily on indirect observation and measurement and use a variety of sources and techniques. The conceptual model has become a *sine qua non* for advance in the field. Physical models of molecules and lattices have proved invaluable. Analysis and synthesis are mutually supportive techniques in the basic problem of elucidation of structure. Uniform systems of nomenclature are vital as part of communication within the discipline. Chemistry helps support and relies on other disciplines. Chemists do not work entirely without responding to societal demands.

Inquiry in education

Inquiry is regarded as an important feature of independent and autonomous learning, which many educators currently regard as desirable outcomes in themselves. Some learning theories support the inquiry mode, e.g. the theory of cognitive dissonance. Inquiry techniques are associated with, but are not to be confused with, problem solving and discovery methods. The use of inquiry needs to be supported by other practices, and inquiry is most effective when students are taught to use the method and have had practice in it. Inquiry requires knowledge on the part of the student before effective inquiry questions can be asked. Individual differences cannot be ignored when considering the use of inquiry techniques. Inquiry techniques include the use of conceptual models.

Inquiry in the teaching of a science

The position taken in this paper, then is that there are two aspects of inquiry. (Schwab, 1962). The complete inquiry classroom would have these two aspects. On the one hand, the student would be led to inquire into these materials. Schwab suggests, therefore, that there is a deliberate ambiguity in the phrase "teaching science as inquiry". Gagné (1963, p. 144) speaks of a widespread consensus that an instructional programme for the science student "most clearly achieves its rightful goal when it enables such a student to adopt the procedures of scientific enquiry in response to any new unsolved problem he encounters". There is the need, then, for scholars in the field to be involved in curriculum to facilitate the attainment of such a goal. There are limits to expectations in this direction. Psychological theories of learning and the preparation of teachers to work in this mode are important considerations. The development of special competencies and attitudes is involved.

CHEM study and inquiry

The following is an extract from a pamphlet prepared to explain CHEM Study (Pimentel & Ridgway, 1972) and may be regarded as a significant position statement which highlights claims for inquiry in the materials.

> . . . the student is engaged continually in the pattern of scientific activity, experimental collection of data, assessment and organization of facts, deduction of unifying principles, and application of these principles in developing expectation (making predictions.) He is given the opportunity for discovery. He is invited to think about what goes on around him. He is encouraged to question dogma, to seek definitive tests thereof, to frame his own hypotheses, and to become aware of uncertainty in how the Universe operates. He is tested for his ability to engage in these activities of science, not for his ability to parrot some facts.

In the Parry and Steiner text, the questions are asked, "What does it mean when someone says he is a chemist? How does he spend his time?" This suggests an attempt in the materials to meet that important current aim in the teaching of science that has been mentioned.

Klopfer's categories

What was required was a set of categories that provides a checklist when used as a coding manual for written chemistry curriculum materials. Klopfer (1971) devised a table of specifications for science education, in which the stated focus is on categories of student behaviour related to the carrying out of the processes of scientific inquiry. Klopfer argues that,

> This focus is justified, not only because the contemporary trend of science education is toward an emphasis in the processes of inquiry, but because science is meaningfully and significantly considered as a system of inquiry, rather than simply as structured knowledge.
>
> (Klopfer, 1971, p. 568)

The categories used in this study, based on Klopfer's analysis, are indicated in Table 1. Ratings were made of the extent to which inquiry was emphasized in each chapter and experiment of the Parry materials on a five point scale according to whether there was no referance made to the category (1), no more than two references (2), three to four references (3), five to six references (4) or more than six references (5).

Validity and reliability checks

The instrument, based on Klopfer's table of specifications, includes categories spanning the construct of inquiry as revealed by an analysis of the literature and was shown to possess adequate content validity for the purposes of this study. Subjectivity in the use of a rating scale always poses a problem of reliability. Teachers from a large private school in Brisbane and a College assisted in attempts to eliminate this problem and the coding manual was used by the author on four separate occasions, with high product-moment correlation between the third and fourth scorings.

TABLE 1

Specifications Table for use as a coding manual: The Instrument.

CHEMISTRY: EXPERIMENTAL FOUNDATIONS TEXT CHAPTER
LABORATORY MANUAL EXPERIMENT

B. PROCESS OF SCIENTIFIC INQUIRY I: OBSERVING AND MEASURING

		1	2	3	4	5
B1	Observation of objects and phomena					
B2	Description of observations using appropriate language					
B3	Measurements of objects and changes					
B4	Selection of appropriate measuring instruments					
B5	Estimation of measurements and recognition of limits in accuracy					

C. PROCESSES OF SCIENTIFIC INQUIRY II: SEEING A PROBLEM AND SEEKING WAYS TO SOLVE IT

		1	2	3	4	5
C1	Recognition of a problem					
C2	Formulation of a working hypothesis					
C3	Selection of suitable tests of a hypothesis					
C4	Design of appropriate procedures for performing experiments					

D. PROCESSES OF SCIENTIFIC INQUIRY III: INTERPRETING DATA AND FORMULATING GENERALIZATIONS

		1	2	3	4	5
D1	Processing of experimental data					
D2	Presentation of data in the form of functional relationships					
D3	Interpretation of experimental data and observations					
D4	Extrapolation and interpolation					
D5	Evaluation of a hypothesis under test in the light of data obtained					
D6	Formulation of generalizations warranted by relationships found					

E. PROCESSES OF SCIENTIFIC INQUIRY IV: BUILDING, TESTING AND REVISING A THEORETICAL MODEL

		1	2	3	4	5
E1	Recognition of the need for a theoretical model					
E2	Formulation of a theoretical model to accommodate knowledge					
E3	Specification of relationships satisfied by a model					
E4	Deduction of new hypotheses from a theoretical model					
E5	Interpretation and evaluation of tests of a model					
E6	Formulation of a revised, refined, or extended model					

Explanation of the Manual for use with CHEM Study Materials.

Listed below are two examples of the behaviour categories for inquiry. For each category, the code used, a brief description of the category (based on Klopfer's analysis) and a selected example from the CHEM Study Text and/or Laboratory Manual to illustrate its use are given. The examples used are all considered examples of inducing inquiry behaviour in students.

Code	*Specification*
B1	*Observation of objects and phenomena.* Observation includes the use of any of the five senses and their extension to instruments. Text: Some (electrical phenomena), like the electric charge we generate when walking across a wool rug in winter, have been known for centuries. Others are comparatively new to us. Try to name five you have seen (p. 119). Manual: For several minutes observe any changes that take place (p. 18).
E6	*Formulation of a revised, refined, or extended model.* As Klopfer (1971, p 575) puts it, "The science student engaging in fluid inquiry will before long encounter a phase of theory building in which he finds it necessary to reformulate a theoretical model he has espoused." The reformulation may range from a minor modification to major surgery. Text: How can the model be expanded to accommodate this new set of observations? (p. 128). Manual: How does a liquid differ from a gas in compressibility? Interpret this fact by a modification of the model used to describe gases. (p. 32).

Results

The data showing the counts of the times the various Klopfer categories occur in the CHEM Study materials is given in the following figures.

Discussion of results

From a study of the data, a number of points may be made:-

(i) The extent of the CHEM Study materials, as a curriculum package, to use and foster the principles of inquiry is most clearly shown in Figure 3 which shows the mean on all criteria for each text chapter supported by experiments from the laboratory manual suggested in the teachers guide; and in Figure 4 which shows the mean by all criteria across all text chapters supported by experiments from the laboratory manual suggested in the teachers guide. The experiments in the laboratory manual support the chapters in using inquiry.

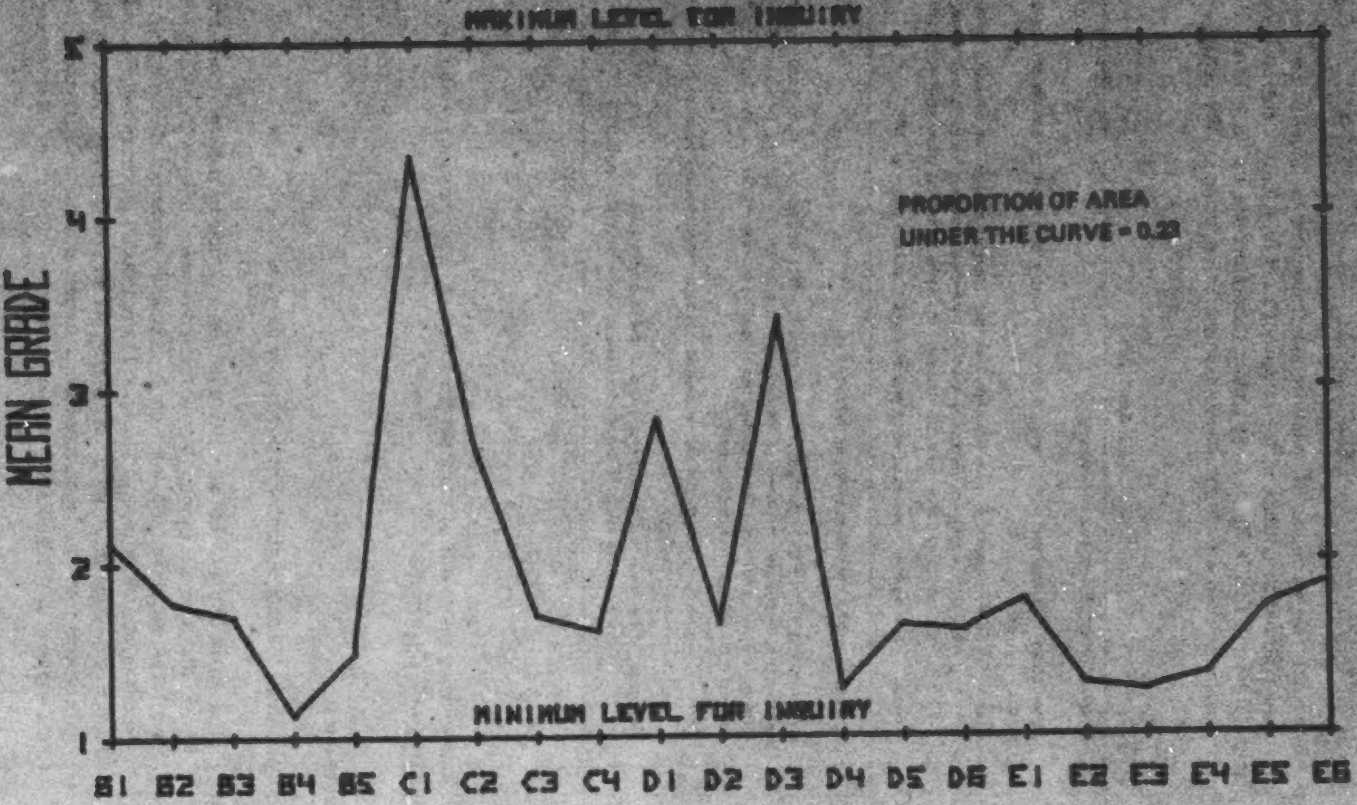


FIGURE 1

Specification for inquiry : Text (all chapters)

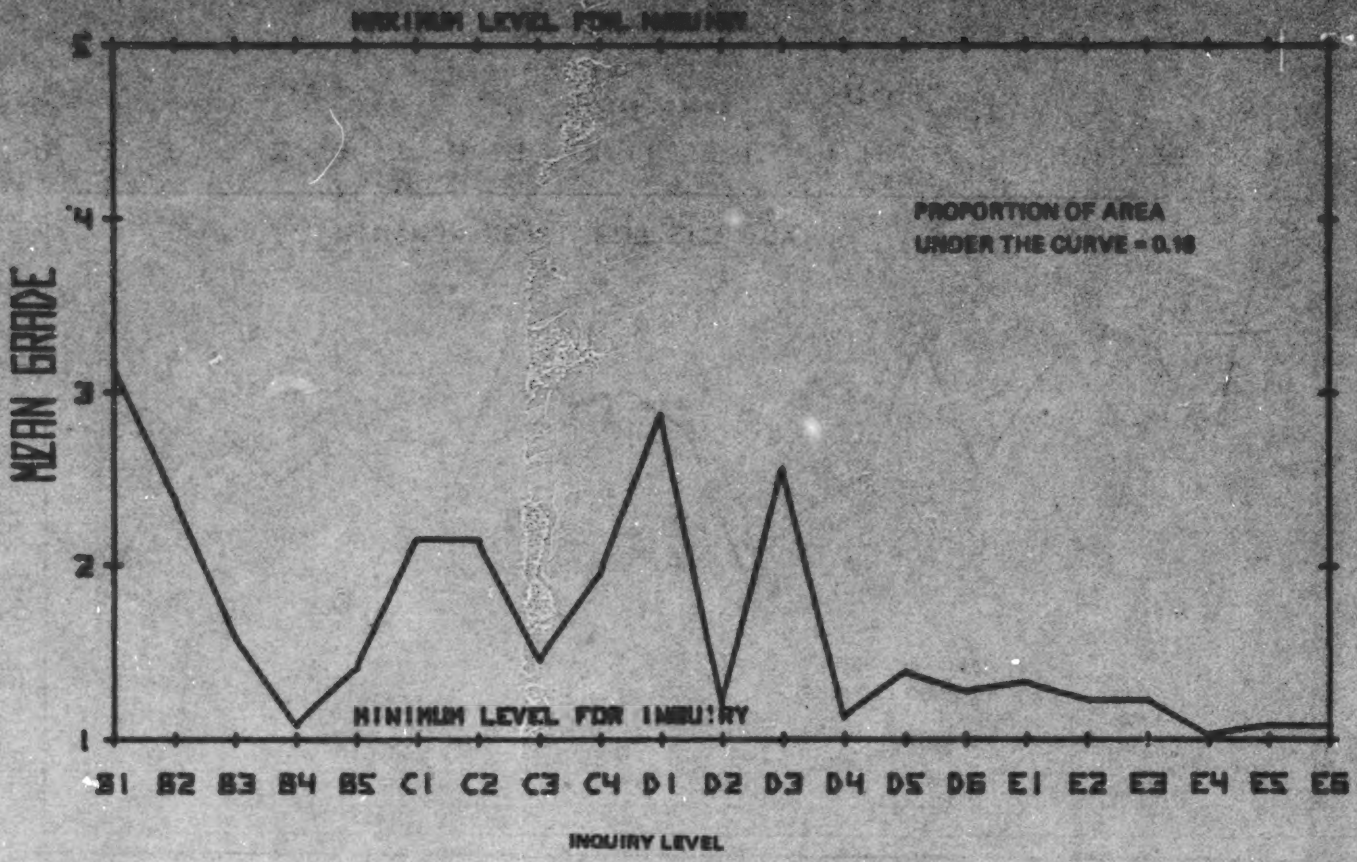


FIGURE 2

Specification for inquiry : Lab. Manual (all experiments)

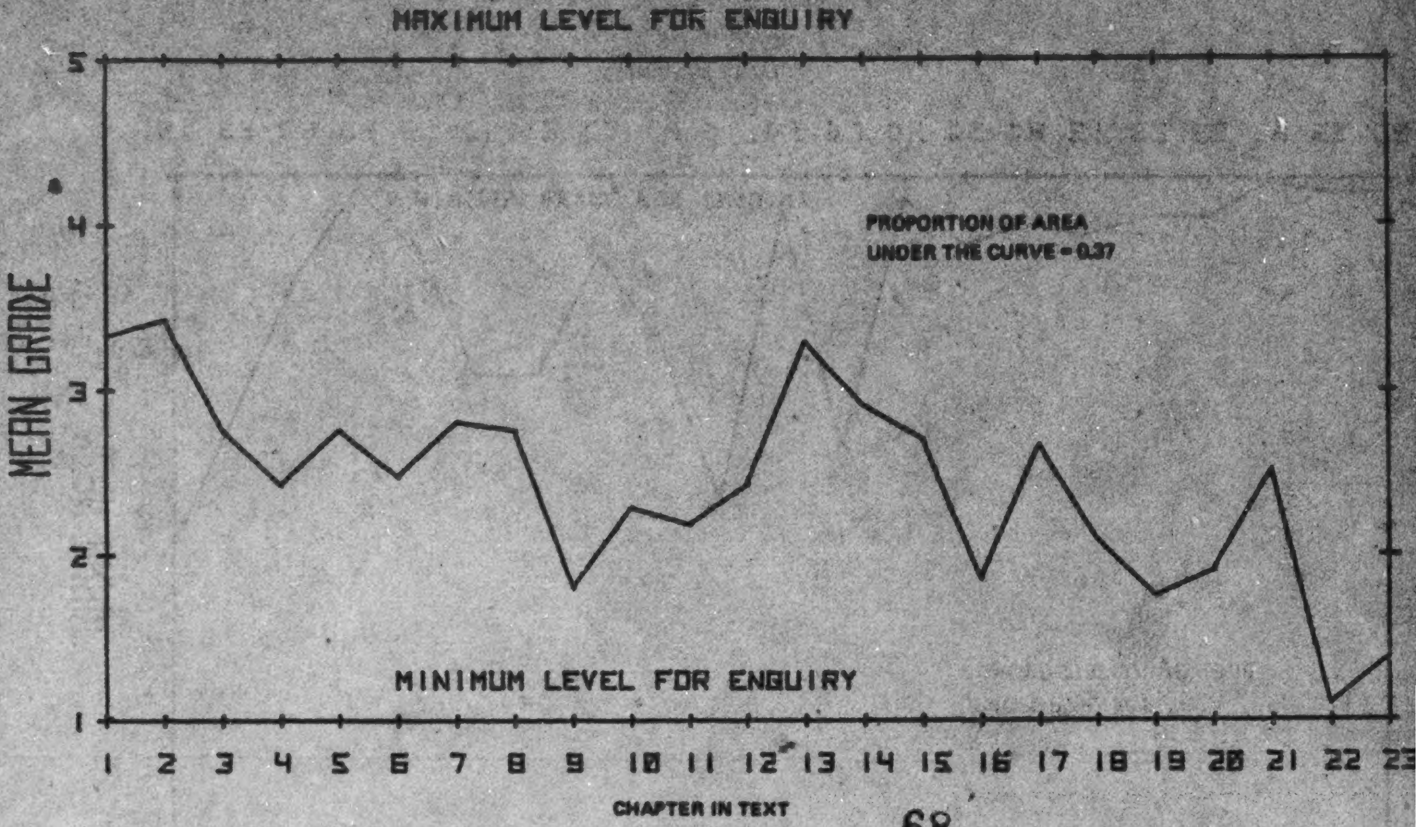


FIGURE 3

Text and laboratory manual

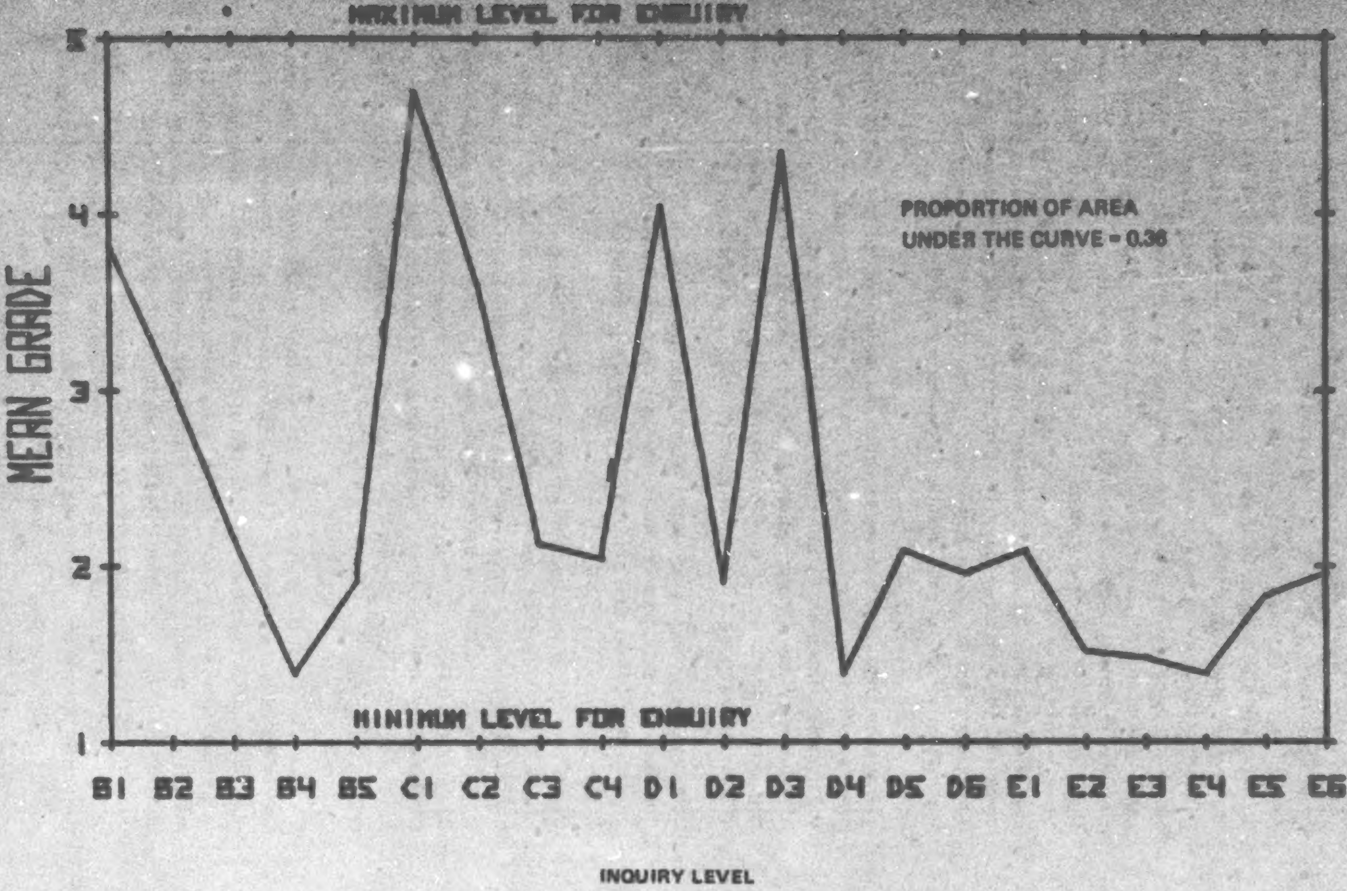


FIGURE 4

Specification for inquiry : Text supported by experiment

(ii) The data represented in Figure 1 (the Text Chapters), Figure 2 (Laboratory Manual) and Figure 3 (Text supported by the Manual) indicate that there is no great consistency in the use of inquiry, with, in fact, quite definite peaks and lows. These same three graphs, and particularly Figure 1, indicate a general steady decline in the use of inquiry. Chapters 1, 2, 4 and 5 of the text have the highest levels of inquiry. In Chapters 1-6 of the text, including support from the laboratory manual, it seems that the main thrust of the writers, as far as the principles of inquiry are concerned, is the development of the concepts in a variety of ways, ways that are consistent with Klopfer's categories. One may reasonably assume that the writers feel confident that in later chapters they can call on this development with a further occasional heavy "injection" towards the middle of the text (Chapters 13, 14, 15). One is left feeling, however, that Chapters 22 and 23 involve "tacked-on" topics as far as inquiry is concerned (if in no other way). It could be argued that concluding sections of curriculum material should come back to re-inforce stated objectives of the material, which, in this case, include the development of inquiry.

(iii) The graphs also suggest some strengths and weaknesses of the CHEM Study material in inquiry. Firstly, categories B4, C3 and C4 are low. In defence of the materials, however, CHEM Study does not claim to cover these categories of inquiry to any serious extent. It could be argued, of course, that such materials in chemistry curriculum should in fact do this, as the analysis of inquiry as a component in chemistry and in education might suggest.

Category D2 is also low for both text and laboratory manual. Klopfer's use of this category seems, however, to be narrowly restrictive, expecting the relationships mentioned here to lead readily to graphical representation.

(iv) On the other hand the interpretation of categroy C1 "Recognition of a Problem" is very broad indeed and scores highly, particularly in the text. This category proved to be one of the more difficult in interpretation and in determining its scope. It was here that the borderline between a simple rhetorical question and the category proper was fuzzy and sections in the materials usually received the benefit of the doubt. This seemed at least reasonably fair, when the style in the CHEM Study materials is compared with chemistry texts of similar level of difficulty.

(v) Concern was felt at the general low level of categories D4 through E6 particularly in the laboratory manual. E1 and E6 are reasonably high in the text. This may be expected since they may be seen as mutually supportive categories. One may reasonably ask if, in fact, Klopfer's scheme expects too much of curriculum materials by spreading the building, testing and revising of a theoretical model over too many (six) categories. Klopfer's position could be defended since model building has been seen as a *sine qua non* for advance in the field of chemistry and so pertinent to inquiry; even if, therefore, these six categories were condensed to four or even three, the total use of them in CHEM Study materials would remain low enough to cause some concern.

(vi) The figures also show how the picture presented in earlier figures is modified when the support given to the text chapters by the experiments in the laboratory manual is considered. Chapters 9, 22 and 23 remain low, chapters 2, 3, 7, 12, 13 are noticeably improved, whereas most of the rest indicate no more than the kind of increase expected by introducing additional material for support in the way of one or two experiments. The analysis does not indicate directly the relevance of the supporting experiments but a reading of the materials suggests that this is acceptably high.

References

BEBELL, C. Evaluation and curriculum development. *Educational Leadership*, 1962, 17, 4-6. Quoted in Payne, D.A., *Curriculum evaluation*. Lexington, Massachusetts: D.C. Heath and Company, 1974, 22-25.

BLOOM, B.S. The role of the educational sciences in curriculum development. *International Journal of Educational Sciences*, 1967, 1, 7. Quoted in M.D. Herron, The nature of science inquiry. *School Review*, 1971, February, 171-213.

CHAMPAGNE, A.B. & KLOPFER, L.E. Formative evaluation in science curriculum development. *Journal of Research in Science Teaching*, 1974, 11(3), 105-203.

GAGNÉ, R.M. The learning requirements for inquiry. *Journal of Research in Science Teaching*, 1963, 1(2), 144-153.

HERRON, M.J. The nature of scientific enquiry. *School Review*, 1971, 79, 171-212.

KLOPFER, L.E. Evaluation of learning in science. In B.S. Bloom, J.T. Hastings & G.P. Madaus, *Handbook on formative and summative evaluation of student learning*. New York: McGraw-Hill, 1971, 559-641.

LEE DOW, K.C. *Teaching science in Australian schools*. Melbourne: Melbourne University Press, 1971.

LUCAS, A.M. Creativity, discovery and inquiry in science education. *Australian Journal of Education*, 1971, 15(2), 185-196.

PARRY, R.W., STEINER, L.E., TELLEFSON, R.L. & DIETZ, P.M. *Chemistry: experimental foundations*. Englewood Cliffs, New Jersey: Prentice-Hall, (1964).
(Also Teachers Guide and Laboratory Manual).

PAYNE, D.A. *Curriculum evaluation: commentaries on purposes, process, product*. Lexington, Massachusetts: D.C. Heath, 1974.

PIMENTAL, G.C. & RIDGWAY, D.M. Chem study: knowledge of chemistry. *Science Activities*, 1972, November, 40-45.

SCHWAB, J.J. The concept of the structure of the discipline. *Education Record*, 1972, July, 197-209.

WOLF, R. A model for curriculum evaluation. *Psychology in Schools*, 1969, 6, 107-108. Reprinted in D.A. Payne (ed.), *Curriculum evaluation: commentaries on purpose, process, product*. Lexington, Massachusetts: D.C. Heath, 1974.

Research in Science Education, 1978, 8, 71-79

THE EFFECTS OF JUNIOR INQUIRY SCIENCE PROGRAMS ON STUDENT COGNITIVE AND ACTIVITY PREFERENCES IN SCIENCE

John Dekkers

Introduction

In order to gain a broader perspective of the influences of new science curriculum materials and approaches, recent researches have paid increasing attention to a number of intricately related student characteristics that influence the learning process. For instance, learners' acquired characteristics may interact with other individual characteristics and thereby influence further outcomes in subsequent learning under specific modes of instruction and other conditions. Attitude (Krathwohl *et al.*, 1964), cognitive style (Glaser, 1970), cognitive structures (Ausubel, 1963) and cognitive preferences (Heath, 1964) are examples of such characteristics.

The acquired student characteristic cognitive preference has been the subject of numerous researches in science education. Most of the studies have evolved from Heath's conceptualisation of cognitive preference, which he defined as "the differing modes of attending to the subject matter of a course" (Heath, 1964, p. 241). The nature of science cognitive preferences and results of cognitive preference testing have been recently examined (Brown, 1975; Tamir, 1975, 1976a; Mackay, 1976) and findings indicate that substantial and consistent agreement exist between results from a number of independent studies. Also, the results intimate the utility of cognitive preference testing in curriculum evaluation.

However, cognitive preference tests in science have tended to focus on a particular science discipline. An exception are the tests designed by Atwood (1964) for general science. Furthermore, cognitive preference testing has tended to be confined to students at upper secondary school or at college level. The present study attempts to assess the impact of inquiry science materials on students in their first year of high school.

This paper introduces the notion of activity preferences in science. An activity preference is a students preferred means for acting upon an observation or investigating a question or event. An instrument developed to measure activity preferences is described below.

The activity preferences exhibited by students in science programs that differ in nature and manner of student experience provided may be expected to differ significantly. Five activity modes for which activity preferences were measured in this paper are: (a) laboratory work; (b) field work; (c) individual student project work; (d) discussion and (e) reading activities. These activities feature most predominantly in recent science curriculum developments that emphasize active student participation in a variety of experiences within a framework of common inquiry goals. The science curricula Australian Science Education Project (ASEP) and Foundational Approaches in Science Teaching (FAST), the subjects of this study, utilize the above activities in biological, earth and physical science studies.

This paper describes a cognitive and an activity preference test and their use with students studying either FAST or ASEP. These curricula were examined because they differ considerably in the organisation of subject matter whilst defined teacher and student roles in each curriculum are very similar. FAST is a three year sequential laboratory and field centred "inquiry" science for junior science students (Pottenger *et al.*, 1974).

> The program consists of three strands; Physical Science, Ecology and Relational Study. The Physical Science and Ecology strands provide studies in several sciences including

> chemistry, physics, biology, geology and meteorology. The integrity of the various disciplines is maintained in these strands. The Relational Study strand provides a way of viewing the interrelationships between the sciences, technology and society.
>
> (Dekkers and Rouse, 1977, p. 82)

In contrast ASEP presents teachers and students with a selection of materials which emphasize the environment of the student (ASEP, 1974). Furthermore, ASEP is not disciplinary oriented.

It was thus interesting to examine the possible effects of different orientations of curriculum organisation on student cognitive and activity preference.

Design of study

The sample Students (N=190) located in two high schools comprised the population. One high school used ASEP (N=88; 53 males, 35 females: 3 classes). The other school used FAST (N=101, all females: 3 classes). All students were administered 3 tests — Science Cognitive Preference Inventory (SCPI), Activity Preference in Science (APIS) and a science achievement test. Testing was administered by the author in August-September, 1976.

Instruments

Science Cognitive Preference Inventory (SCPI)

The test contains 30 items of which 18 items relate to biological sciences and 12 items to physical science. The four cognitive areas originally suggested by Heath (1964) are utilized. The test is designed to assess the relative preferences, towards the following four areas: recall of facts (R); principles and explanations (P); critical questioning and demand for more information (Q) and practical application (A). SCPI is a modified version of a 40 item cognitive test developed by Tamir *et al*. (1976b).

The key for each item, that is the allocation of the four cognitive preference areas, was determined by 8 expert judges and only items reaching at least 90 per cent agreement were included.

In the selection of responses, students were requested to arrange the options within an item in order of preference by allocating 1 for the most preferred option, 2 for the next preferred option, 3 for the next option, and 4 for the least preferred option.

The data was so analysed that all responses appearing in an item contributed to a students overall cognitive preference profile. The lowest possible score for any preference area is 30 and highest is 120. A complete response yields a total score of 300. A random response distribution yields a mean score of 75 for each of the cognitive preference areas. In this paper all scores are converted to preference area item means.

A test-retest reliability carried out on Grade 8 high school students (N=59; 33 males, 26 females) gave the following correlations:

τ (recall) = 0.81; τ (principle) = 0.65; τ (questioning) = 0.84; τ (application) = 0.79.

Activity Preference in Science (APIS) (Dekkers, 1976)

The test, consists of 30 items of which 17 items relate to biological science and 13 items to physical science. It seeks to elicit the extent of student preference for the school activities laboratory work, discussion, field work, reading, and performing individual projects. A typical item is:

> A certain insect survived in a refrigerator for 10 hours. Activities available to gain more information are:
> (A) Read about how insects manage to survive at cold temperatures.
> (B) Devise experiments to determine the best survival conditions for the insect at cold temperatures.
> (C) Have a class discussion about the possible adaptations that may be used by the insect to survive in cold temperatures.

Option (A) corresponds to a reading activity, (B) devising experiments and (C) a discussion activity. Each item always contained three different types of activity responses. The total test was so devised that each different activity response was contained in a minimum of 16 items.

The key was determined by expert judges in the same manner as for SCPI.

In the selection of responses, students were requested to arrange the options within an item in order of preferences by allocating 1 for the most preferred activity, 2 for the next preferred activity, and 3 for the least preferred activity.

Like the SCPI, all student responses contributed to a student's overall activity preference profile. The lowest possible score is 16 and highest is 64. A complete response yields a total score of 180. A random distribution yields a mean score of 32 for each of the cognitive areas. In this paper all scores are converted to preference area item means.

A test-retest reliability carried out on Grade 8 high school students (N=46; 29 males, 17 females) gave the following correlations:

$$\tau\ (\text{lab.}) = 0.85;\ \tau\ (\text{discussion}) = 0.72;\ \tau\ (\text{field work}) = 0.69;\ \tau\ (\text{reading}) = 0.87;$$
$$\tau\ (\text{project}) = 0.62.$$

Achievement Test

The test was specifically designed for the study. It consisted of 30 multiple choice item questions and short answer questions based on content familar to all students.

Results and discussion

Science Cognitive Preference Inventory

Mean scores, reliability coefficients and F test results for the total sample of FAST and ASEP students are presented in Table 1.

In examining these results it should be noted that in the Student Response Sheet for SCPI the most preferred cognitive mode is given 1 by the student, the next preferred is given 2, the next 3 and the least preferred mode 4. Thus high preferences are indicated by low means.

The preferred cognitive modes for students in the FAST and ASEP curriculum have been ordered in Table 3, based on mean scores and correlated t-values for cognitive modes (Table 2).

TABLE 1

Mean cognitive preference scores, F values and reliability coefficients

Cognitive Preference Area	*FAST Curriculum N=104* $\bar{X}$	*SD*	*Cronbach α Coefficient*	*ASEP Curriculum N=87* $\bar{X}$	*SD*	*Cronbach α Coefficient*	*F (df=1, 189)*
All Items							
Recall	2.48	0.36	0.62	2.59	0.30	0.56	4.1**
Principles	2.57	0.25	0.38	2.54	0.22	0.30	NS
Questioning	2.44	0.37	0.64	2.43	0.35	0.66	NS
Application	2.51	0.25	0.42	2.43	0.23	0.36	NS
Biology Items							
Recall	2.45	0.41	0.51	2.61	0.32	0.46	7.1*
Principles	2.65	0.29	0.39	2.54	0.24	0.33	4.0*
Questioning	2.49	0.42	0.53	2.44	0.41	0.59	NS
Application	2.41	0.35	0.38	2.41	0.25	0.29	NS
Phys. Sc. Items							
Recall	2.52	0.46	0.50	2.57	0.39	0.39	NS
Principles	2.45	0.38	0.38	2.55	0.31	0.26	4.2**
Questioning	2.36	0.46	0.48	2.41	0.41	0.45	NS
Application	2.66	0.33	0.41	2.46	0.36	0.32	8.3*

Only significant F values are reported **p = 0.05, *p = 0.01 level.

TABLE 2

Correlated t-values[1] for cognitive modes

Cognitive Mode Pairs	*FAST Curriculum N=101* *All Items*	*Biology Items*	*Phys.Sc. Items*	*ASEP Curriculum N=87* *All Items*	*Biology Items*	*Phys.Sc. Items*
R:P	NS	3.7*	NS	NS	NS	NS
R:Q	NS	NS	NS	3.2*	3.5*	3.5*
R:A	NS	NS	2.3**	3.5*	4.1*	3.8*
P:Q	3.4*	3.6*	NS	2.3**	2.4**	2.7*
P:A	NS	4.5*	3.7*	2.3**	2.9*	2.4**
Q:A	2.8*	NS	4.5*	NS	NS	NS

[1] Only statistically significant t-values are reported.

**p = 0.05

*p = 0.01 level for two tailed probability.

TABLE 3

FAST Curriculum	*ASEP Curriculum*
All Items	
$Q \approx R > A \approx P$	$Q \approx A > P \approx R$
(most preferred – least preferred)	(most preferred – least preferred)
Biology Items	
$A \approx R \approx Q > P$	$A \approx Q > P \approx R$
Physical Science Items	
$Q \approx P \approx R > A$	$Q \approx A > P \approx R$

The results in Table 3 reveal that both FAST and ASEP students have high preferences for the Questioning mode. Since the FAST and ASEP curricula both aim to promote in students a sense of objective criticism, open-mindedness and curiosity, a high preference for Q may be regarded as a favourable outcome regarding the achievement of these objectives.

The results of the subtests reveal the essential differences between FAST and ASEP students. However, Table 1 reveals that the reliability coefficients of the subtest results for both FAST and ASEP are generally low and thus the results are being cautiously interpreted. For ASEP students cognitive modes have grouped into pairs Questioning/Application and Principles/Recall for the total test and subtests. There being a strong preference for Q/A and a low preference for P/R. For FAST students a different pattern emerges. It would appear that the cognitive preferences for FAST students are to some extent dependent on specific subject matter. Application is the most preferred mode in the biological sciences and least preferred in the physical sciences. Presumably in the FAST program the usefulness and applicability of biology is considered by students to be more important than in physical science. It is suggested that the above difference may be attributable to the structural difference between FAST and ASEP as outlined in the Introduction.

Student results of the science achievement test were divided into four quarters. Analysis of variance indicated no significant interactions between the achievement categories for each curriculum. Further, no significant interactions were observed between sex and the curriculum studied. Results of cognitive preference tests in studies using senior high school and college students indicate that cognitive preferences are related to achievement and sex (Mackay, 1975). However a recent study by Dekkers and Allen (1976) which examined student cognitive preferences in junior high schools revealed that cognitive preferences were independent of achievement and sex for students in the first year of high school but were dependent on achievement and sex in the later years of high schools.

It should be noted that the Science Cognitive Preference Inventory scale is relative rather than absolute. Thus a low preference does not necessarily imply a dislike it could be that a student likes other cognitive modes better. Similarly, the highest preference may not mean a real liking.

Activity Preference in Science

Mean scores, reliability coefficients and F test results for the total sample of FAST and ASEP students are presented in Table 4.

TABLE 4

Mean activity preference scores, F values and reliability coefficients

Activity Preference Area	*FAST Curriculum N=103* $\bar{x}$	*SD*	*Cronbach α Coefficient*	*ASEP Curriculum N=87* $\bar{x}$	*SD*	*Cronbach α Coefficient*	*F (df=1, 187)*
All Items							
Laboratory work	1.89	0.24	0.62	1.79	0.27	0.58	7.5*
Discussion	2.15	0.27	0.78	2.24	0.40	0.83	NS
Field work	1.54	0.36	0.69	1.81	0.34	0.63	4.0*
Reading	2.37	0.34	0.77	2.10	0.38	0.78	8.9*
Project work	2.13	0.40	0.82	1.96	0.36	0.74	4.4**
Biology Items							
Laboratory work	1.85	0.27	0.41	1.86	0.33	0.46	8.3*
Discussion	2.70	0.37	0.64	2.26	0.43	0.72	NS
Field work	1.54	0.38	0.54	1.74	0.35	0.43	5.9*
Reading	2.38	0.37	0.55	2.08	0.37	0.56	6.4*
Project work	2.16	0.43	0.62	1.96	0.40	0.52	4.0**
Phys. Sc. Items							
Laboratory work	1.50	0.37	0.54	1.69	0.36	0.47	5.1**
Discussion	2.10	0.39	0.64	2.23	0.44	0.70	NS
Field work	1.50	0.42	0.47	1.90	0.45	0.43	12.1*
Reading	2.37	0.37	0.70	2.11	0.42	0.69	8.4*
Project work	2.10	0.42	0.70	1.94	0.38	0.63	6.1*

Only significant F values are reported **p = 0.05, *p = 0.01 level.

TABLE 5

Correlated t-values[1] for activity modes

Activity Mode Pairs	*FAST Curriculum N=101* *All Items*	*Biology Items*	*Phys.Sc. Items*	*ASEP Curriculum N=87* *All Items*	*Biology Items*	*Phys.Sc. Items*
L:D	10.7*	7.8*	11.9*	10.7*	5.6*	9.8*
L:F	2.4**	2.8**	NS	NS	2.4**	3.2*
L:R	13.8	12.2*	23.7*	10.4*	4.2*	3.8*
L:P	10.7*	9.8*	18.6*	13.2**	3.6*	4.3*
D:F	11.0*	10.1*	15.9*	5.6*	7.2*	5.9*
D:R	4.0*	3.5*	4.7*	NS	3.8*	3.4*
D:P	NS	NS	NS	2.3**	5.8*	5.9*
F:R	14.4*	12.8*	21.8*	4.5*	5.3*	7.7*
F:P	11.1*	13.2*	15.9*	2.3**	3.7*	NS
R:P	4.7*	3.9*	4.0*	3.2*	NS	3.2*

[1] Only statistically significant t-values are reported.

**p = 0.05, *p = 0.01 level for two tailed probability.

In the Student Response Sheet for APIS the most preferred mode is given 1 by the student, the next preferred is given 2 and the least preferred is given 3. Thus in examining these results as with SCPI a high preference for an activity mode is indicated by a low mean. Also the *Activity Preference in Science* scales are relative rather than absolute.

TABLE 6

FAST Curriculum	*ASEP Curriculum*
All items	
$F > L > P \approx D > R$	$F = L > P > R \approx D$
(most preferred – least preferred)	(most preferred – least preferred)
Biology items	
$F > L > P \approx D > R$	$F > L > P = R > D$
Physical Science Items	
$L \approx F > P \approx D > R$	$L > F \approx P > R > D$

The results in Table 6 show that for both FAST and ASEP students that the activity modes tend to group into two sets – Laboratory/Field Work activities and Project/Reading/Discussion activities. L/F activities are most preferred whilst the P/R/D activities are preferred to a significantly lesser extent.

The high preference for F/L together with the high preference for Questioning suggest that students favour a student "inquiry orientated" approach for learning science. The lower preference for R/D may be due to a number of factors that are inherent and common to both curricula. For instance classroom observations of FAST and ASEP classes (Dekkers, 1978) indicate that students are given little encouragement and opportunity to do reading in science. Also much classroom discussion remains teacher centred.

Table 4 reveals significant interactions between activity modes with FAST and ASEP students. These differences largely reflect the extent of preference for a particular science activity. It appears that:

1. FAST students have a greater preference for Laboratory and Field work than ASEP students.
2. FAST students not only have Reading as the least preferred activity mode but also it is a significantly less preferred activity mode compared to ASEP students.

It is suggested that the above differences may be attributable to the structural differences between FAST and ASEP as outlined in the Introduction.

Summary and concluding comment

The preceding findings show that the sample of Grade 8 science students using either FAST or ASEP material have high preferences for questioning, field and laboratory work. This result may be interpreted as a favourable student reaction to learning science by inquiry as intimated by these curricula. It has also been inferred from differences obtained in the ordering of preferences and the extent of differences between preferred preference areas may be due to influences resulting from differences between the structural organisation of these curricula. More in depth investigations are deemed necessary to support this inference.

It further seems necessary to know the stimuli associated with student choice for a particular cognitive or activity preference in science. To investigate this aspect the author is currently investigating the relationship between student cognitive and activity preferences and attitudes and the nature of science teaching practices in junior science classes.

References

ATWOOD, R.K. Development of cognitive preference examination using general science and social science content. *Journal of Research in Science Teaching*, 1967-8, 5, 154-159.

AUSTRALIAN SCIENCE EDUCATION PROJECT (ASEP), *A Guide to ASEP*, (ASEP 401). Melbourne: Government Printer, State of Victoria, 1974.

AUSUBEL, D.P. *The psychology of meaningful verbal learning.* New York: Grune & Stratton, 1963.

BROWN, S.A. Cognitive preference in science: their nature and analysis. *Studies in Science Education*, 1975, 2, 43-65.

DEKKERS, J. *Activity preferences in science - an evaluation instrument for student activity preferences in science.* Hawaii: Curriculum Research and Development Group, University of Hawaii, 1976.

DEKKERS, J. Science practices in junior science classes using 'inquiry science programs'. Paper in preparation, 1978.

DEKKERS, J. & ALLEN, L. Malaysian science students cognitive & activity preferences in science. *Malaysian Sc. Teachers Journal*, 1978 (in press).

DEKKERS, J. & ROUSE, F. Foundational Approaches in Science Teaching (FAST) – a structured 'inquiry' oriented junior science curriculum. *The Australian Science Teachers Journal*, 1977, 23, 81-90.

GLASER, R. Discussion of Messicks Paper, in M.C. Wittrock & D.E. Wiley (ed.), *The evaluation of instruction: issues & problems.* New York: Holt, Rinehart & Winston, 1970.

HEATH, R.W. Curriculum cognition & educational measurement. *Educational & Psychological Measurement*, 1964, 24, 239-253.

KRATHWOHL, D.R., BLOOM, B.S. & MASIA, B.B. *Taxonomy of objectives. The classification of educational goals, handbook II: affective domain.* New York: Mackay, 1964.

MACKAY, L.F. Cognitive preferences & achievement in physics, chemistry, science & mathematics. *Research in Science Education*, 1975, 5, 49-58.

POTTENGER, F., DEMACHE, E. & BRANTLEY, L.R. *Foundational Approaches in Science Teaching (FAST).* Hawaii: Curriculum Research & Development Group, Uni. of Hawaii, 1974.

TAMIR, P. The relationship between cognitive preference, achievement, teacher orientation, curriculum and subject matter. Paper presented at the AERA Annual Meeting, Washington, D.C., 1975.

TAMIR, P. The relationship among cognitive preference, school environment, teachers' curricular bias, curriculum, & subject matter. *American Educational Research Journal*, 1976a, 12, 235-264.

TAMIR, P., *et al. Science Cognitive Preference Inventory (SCPI).* Hawaii: Curriculum Research & Development Group, Uni. of Hawaii, 1976b.

Research in Science Education, 1978, 8, 79-98

MEASURING PERCEPTIONS OF INQUIRY AND OPEN LEARNING ENVIRONMENTS

A. John Rentoul and Barry J. Fraser

Introduction

In recent times, there has been a considerable upsurge in interest among educational researchers in the study of classroom learning environment. This interest is highlighted in two simultaneous but independent research programmes, one instigated by Herbert Walberg as part of the research and evaluation activities of Harvard Project Physics, and the other instigated by Rudolf Moos at Stanford University. The instrument used in Walberg's research was the *Learning Environment Inventory* (LEI), which measures students' perceptions of psychosocial aspects of the classroom like diversity, competitiveness and formality (Anderson, 1973), while the instrument used in Moos' research was the *Classroom Environment Scale* (CES), which measures psychosocial aspects like involvement, rule clarity and innovation (Moos & Trickett, 1974).

The educational literature abounds with articles reporting the use of Walberg's LEI in research, including science education research (see reviews of Anderson & Walberg, 1974; Walberg, 1976, 1978). Moreover several Australian studies have involved the use of modified versions of the LEI in research related to the Australian Science Education Project (ASEP) (Tisher & Power, 1975; Fraser, 1976; Northfield, 1976). In contrast to the large number of articles reporting use of the LEI, the educational literature is almost completely devoid of papers reporting the use of Moos' CES. In fact the present authors are aware of no study in science education which has employed the CES. Unfamiliarity with Moos' work among educational researches can be traced to the fact that major studies of classroom environment have been reported in the psychological rather than the educational literature (e.g. Trickett & Moos, 1973, 1974) and that Moos' work on classroom environment grew out of research in other settings including psychiatric wards (Moos & Houts, 1968), correctional institutions (Moos, 1968), family environments (Moos & Moos, 1976) and community environments (Moos, 1975). Because of the relative unfamiliarity with Moos' work in educational circles, a subsidiary aim of this paper is to bring Moos' writings to the attention of science education researchers (see especially Moos, 1973, 1974; Moos & Insel, 1974; Insel & Moos, 1974).

Past research involving the LEI

Past research involving the use of LEI in school classrooms, including science classrooms, has revealed two consistent patterns of findings (see Walberg, Singh & Rasher, 1977):

1. Student perceptions of learning environment display criterion validity in that they differentiate between classrooms following different curricula or methods of instruction.
2. Student perceptions of learning environment display incremental predictive validity in that they account for appreciable amounts of variance in cognitive and affective learning outcomes beyond that attributable to initial student characteristics.

It is interesting to note that these two patterns of findings were replicated in a study (Fraser, 1978a) involving the use of a modified version of the LEI in junior high school classrooms using ASEP materials (i.e. ASEP classes were perceived quite differently from conventional classrooms, and student perceptions were good predictors of cognitive and attitudinal learning).

In spite of the proven general usefulness of research instruments like the LEI for measuring classroom learning environment, these instruments have important limitations when used in some contexts. First since the LEI was originally developed for use at the senior high school level, readability and suitability for the junior high school cannot be assumed. In fact Power and Tisher (1975) reported that Australian junior high school students did have difficulty in understanding numerous items in the original LEI. Second since the LEI was designed for use in conventional classrooms and was validated with students in these classrooms, it excludes some aspects of learning environment which are particularly relevant in informal and individualised settings. Moreover Power and Tisher (1976) have decried the fact that the vast majority of studies of classroom behaviour have been conducted in a more formal setting in which the teacher is the dominant actor. In commenting on the necessity to devise an instrument for the specific purpose of analysing behaviour in inquiry classrooms they point out that

> . . . a great many systems have been developed for classifying the behaviour of teachers and pupils in teacher-centred classrooms. In general these appear to have limitations when one attempts to use them to describe the many and varied interactions between teachers, pupils and materials in inquiry-oriented, self-paced programs such as ASEP.
>
> (Power & Tisher, 1976, p. 401)

Because of the limitations outlined above, the original LEI was modified by Fraser and Northfield to enhance its suitability for use in individualised junior high school classrooms (see Fraser, 1976, 1978b; Fraser & Northfield, 1976). Modifications made included rewording to improve readability, omission of certain of the original scales which are inapplicable in individualised classrooms, and the development of a new scale called Individualisation to measure the extent to which students perceive their classrooms as individualised and self-paced.

While Fraser and Northfield's single Individualisation scale captured some important aspects of individualised science classrooms, there still remained a need for a new instrument providing a more comprehensive coverage of the salient characteristics of these settings. For this reason a new instrument called the *Individualised Classroom Environment Questionnaire* (ICEQ) has been developed to measure perceptions of those learning environment dimensions which differentiate conventional classrooms from inquiry or open classrooms emphasising individualisation.[1] The main purpose of the remainder of this paper is to describe details of the development and validation of the ICEQ.

Strategy used in developing the ICEQ

The basic strategy adopted for the development of the ICEQ was to identify aspects of the psychosocial environment of the classroom that would simultaneously satisfy the following three criteria:

[1] The ICEQ, which measures perceptions of psychosocial characteristics of individualised classrooms, can be contrasted with Tisher and Power's (1975, 1976) *Scheme for Analysing Behaviour in Individualised Classrooms* (SABIC) for the direct observation and classification of events in individualised classrooms.

1. Dimensions should characterise the learning environment advocated for recent inquiry-based science and social science curricula, or should be frequently mentioned in the literature describing inquiry-based learning and open education.
2. Dimensions should be consistent with the framework delineated by Moos (1973) for conceptualising human environments.
3. Dimensions should be consistent with the opinions and beliefs about inquiry-based and open learning environments held by experts in the field, and should be salient for both practising teachers and school students.

Inquiry-based and open learning environments

Descriptions of learning environments associated with inquiry-based curriculum materials and open education programmes frequently refer to very similar characteristics. The characteristics of open classrooms described by Bussis and Chittenden (1970), Walberg & Thomas (1972) and Traub, Weiss, Fisher & Musella (1972) are similar in many respects to properties of inquiry-based classrooms which have been described by Massialas and Cox (1966), Hurd (1969) and Elliott and Adelman (1975).

In developing DISC (Dimensions of Schooling), an instrument designed to elicit information about the learning environment in open education settings, Traub *et al.* (1972) adapted several of Barth's (1968) key assumptions about the way children behave, develop and learn. Although Barth has noted that these assumptions are hunches based largely on impressions, emotional responses and observations in classrooms, Traub *et al.* considered them to be applicable to open education programmes and to provide a useful starting point for the design of an instrument for describing open classrooms. The learning environment in open classrooms is characterised by Traub *et al.* as involving individualised objectives and activities, student participation in the setting of objectives, a diversity of materials and activities, decentralised decision making, individual or small group work, encouragement of student self discipline and a variety of evaluation methods.

Elliott and Adelman (1975) have identified three attributes of the learning environment which seem to be characteristic of informal, individualised classrooms, whether inquiry-based or open. First the situation is informal – student independence is fostered. Second aims are unstructured - there is considerable flexibility and freedom in regard to intended learning outcomes and procedures, and there is an emphasis on the processes of learning and on individualising the learning environment. Third procedures are guided or open-ended – teaching methods range from responsive guidance with emphasis on means rather than ends, to an open-ended situation which encourages freedom from constraints and maximum student involvement.

The foregoing descriptions of the inquiry-based learning environment favoured in recent science and social science curriculum developments and in the literature concerned with inquiry-based and open education have suggested common constructs for conceptualising the learning environments associated with inquiry-based and open classrooms. Clearly there is a strong emphasis on affective outcomes and the development of student independence, autonomy and responsibility. Usually learning activities and evaluation procedures are individualised, with stress on fostering student interests and promoting attitudes which are important for the personal and social growth of the individual. Students are constantly encouraged to participate rather than to be passive, and the development of skills and processes of inquiry are facilitated by involving students in meaningful investigations and creative activities. Throughout the learning process in these environments, the teacher is very actively involved in the roles of guide and facilitator, counsellor and resource person, adviser and provider, and coordinator and stimulator.

In addition to information from the literature describing inquiry-based and open learning environments, ideas for the development of the ICEQ were also derived from a number of scales, questionnaires, inventories and other systems for measuring or categorising classroom behaviour. Table 1 includes a list of some of these instruments together with salient scales from each instrument. Other useful sources consulted when developing the ICEQ included Adams (1970), Lazarowitz and Lee (1976), Hoyle (1973), Marsh (1976) and Johnson and Johnson (1975).

TABLE 1

Some instruments for assessing classroom learning environment

Instrument	*Authors*	*Examples of Scales*
Classroom Activities Questionnaire (CAQ)	Steele, House & Kerins (1971)	Independent inquiry Participation Concern
Classroom Environment Scale (CES)	Moos & Trickett (1974)	Involvement Affiliation Teacher support
Learning Environment Inventory (LEI)	Anderson & Walberg (1974)	Diversity Formality Competetiveness
Modified LEI	Fraser & Northfield (1975)	Individualization
Open Education Observation Scale	Walberg & Thomas (1972)	Humaneness Instruction Self-perception
Dimensions of Schooling (DISC)	Traub *et al*. (1972)	Setting of instructional objectives Individualization of learning Decision-making
Class Atmosphere Scale (CAS)	Silbergeld *et al*. (1976)	Involvement Autonomy

Moos' classification of human environments

A noteworthy attempt to provide a theoretical framework for studying environments, and one which bears directly on the development of the ICEQ, has been described by Moos (1973). From an intensive study of eight different kinds of environments (broadly classified into treatment environments, total institutions, educational environments and community environments), Moos isolated certain common dimensions involving the following three broad categories:

1. *Relationship dimensions*, which are concerned with the nature and intensity of personal relationships and the extent to which individuals are involved and support each other
2. *Personal development on goal orientation dimensions*, which include variables relating to specific functions of environments and to the potential for peronsal growth and development
3. *System maintenance and change dimensions*, which are concerned with the extent of order and control in the environment and its responsiveness to change.

Insel & Moos (1974) have argued convincingly that these three basic categories can be used as a framework for characterising the broad range of environments mentioned above, and appear to be consistent with scales developed by other investigators.

Development of the ICEQ

In the first stage in the development of the ICEQ, a list of attributes characterising inquiry-based or open classrooms was culled from the literature and from recent curriculum materials as described above. Second, these attributes were grouped together to form intuitively plausible dimensions which were also consistent with the three basic types of dimensions proposed by Moos for characterising human environments. Third, items measuring each dimension were written, using reactions sought from selected experts around Australia, practising teachers and groups of junior high school students. Fourth, the scales were further refined in the light of data collected from field trials of the instrument with several different samples of subjects.

The final version of the ICEQ, after refinement in the light of analysis of data from field trials, consisted of 50 items, with 10 of these measuring each of the following five scales: personalisation, participation, independence, investigation and differentiation.

An important feature of the ICEQ is that, whereas instruments like the LEI measure only students' perceptions of their actual learning environment, the new scales have been designed to measure all four of the following:

1. students' perceptions of their actual classroom learning environment
2. students' perceptions of their preferred classroom learning environment
3. teachers' perceptions of their actual classroom learning environment
4. teachers' perceptions [illegible] their preferred classroom learning environment.

The advantage of having these different formats for the instrument is that it enables the investigation of numerous research questions such as the congruence between teachers' and students' perceptions of the ideal or actual learning environment, and the determinants and effects of discrepancies either between teachers' preferred and actual environment or between students' preferred and actual environment. Furthermore, in writing items for the ICEQ, particular care had to be taken to ensure that the instrument was readily readable by junior high school students and, at the same time, avoided language which "talked down" to teachers.

Each item in the ICEQ is responded to on a five-point scale, ranging from "almost never" to "very often", according to how often the respondent considers that a specific practice either ideally or actually occurs in a classroom. In fact important merits of the ICEQ are that all five scales can be conveniently administered to whole classes of junior high school students in about 20 minutes, and that responses are amenable to objective scoring.

Table 2 provides a description of each scale in the ICEQ together with a sample item from each scale. Also the table provides general descriptions and specific examples of the basic dimensions of human environments proposed in Moos' scheme, and shows how each dimension in the ICEQ can be classified according to Moos' scheme.

Validation of the ICEQ

A preliminary version of the ICEQ was refined to form the final version containing 10 items per scale after analysis of data collected from trial administration to two separate samples. The first sample consisted of 230 eighth and ninth grade students, in eight representative schools in the Sydney metropolitan area, who provided their perceptions of their actual classroom learning environments. The second sample consisted of 120 Sydney secondary and primary teachers, both experienced and in training, who provided their perceptions of their preferred classroom learning environment.

Data from these two administrations were analysed according to the techniques advocated by Fraser (1977) in order to identify items whose removal would improve scale internal consistency (the extent to which a given scale measures the same dimension) and discriminant validity (the extent to which a given scale measures a unique dimension not measured by other scales in the instrument). While scale internal consistency was enhanced by removing any item with a low correlation with the total for its scale, scale discriminant validity was enhanced by removing any item whose correlation with its a priori assigned scale was smaller than its correlation with any of the other four scales.

After the item analysis techniques described above had been applied to reduce each original scale to 10 items, α reliability coefficients were calculated as indices of the internal consistency of the refined scales and scale intercorrelations were computed as indices of discriminant validity. Moreover, because improvements in scale characteristics obtained by removal of certain items are likely to be lost in subsequent administrations to other samples because of sampling variations (Ebel, 1961), the refined scales were crossvalidated using a third sample to check the scale statistics. This third sample consisted of 81 trainee secondary teachers in Sydney who provided their perceptions of their preferred classroom learning environments.

Table 3 shows the values obtained at each administration for scale internal consistency (α coefficients) and discriminant validity (scale intercorrelations). This table indicates that α coefficients were all satisfactorily high and ranged from 0.65 to 0.81 for the sample of students, from 0.72 to 0.86 for the first sample of teachers, and from 0.62 to 0.81 for the crossvalidation sample of teachers. The intercorrelation matrices in Table 2 also suggest that the scales displayed satisfactory discriminant validity since scale intercorrelations were generally moderately low for all three administrations and the mean correlation over all administrations was less than 0.3.

Conclusion

This paper can be summarised conveniently by listing the following positive features of the *Individualised Classroom Environment Questionnaire*: its dimensions are consistent

TABLE 2

Description of scales in the Individualised Classroom Environment Questionnaire (ICEQ) and their classification according to Moos' scheme for conceptualising human environments

Individualised Classroom Environment Questionnaire (ICEQ)			*Moos' Classification of Human Environments*		
Scale	*Description of scale*	*Sample item*	*Basic type of dimension*	*Description of basic type of dimension*	*Examples of specific dimensions*
Personalisation	Opportunities are provided for individual students to interact with the teacher and there is concern for the personal welfare and social growth of the individual	The teacher takes a personal interest in each student.	Relationship	The nature and intensity of personal relationships within the environment. The extent to which individuals are involved in the environment and the extent to which they support and help each other.	Involvement Affiliation Staff support Peer cohesion
Participation	Students are encouraged to participate rather than be passive listeners.	Students discuss their work in class.			
Independence	Students are allowed to make decisions about, and to have considerable control over, their own learning and behaviour.	Students choose their own partners for group work.	Personal Development or Goal Orientation	The potential or opportunity in the environment for personal growth and the development of self-esteem. Goal orientation variables relate to specific functions of the environment, e.g. "academic style" in the classroom.	Academic achievement Practical orientation Competition Task orientation
Investigation	There is emphasis on the skills and processes of inquiry and their use in problem solving and investigation.	Students draw conclusions from information			
Differentiation	There is emphasis on the selective treatment of students on the basis of ability, learning style, interests, and rate of working.	Students work at their own speed.	System Maintenance and System Change	Extent to which the environment is orderly and clear in its expectations, maintains control, and is responsive to change.	Order and organisation Clarity Control Innovation

TABLE 3

Internal consistency (α reliability) and discriminant validity (scale intercorrelations) for three administrations of the ICEQ

Perceptions Measured	*Sample Size*	*Scale*	*α Reliability*	*Scale Intercorrelations*				
				Pers	*Part*	*Ind*	*Inv*	*Diff*
Students' Actual Environment	230	Personalisation	0.81	-	0.58	0.32	0.48	0.02
		Participation	0.70		-	0.22	0.41	0.07
		Independence	0.70			-	0.03	0.33
		Investigation	0.85				-	0.12
		Differentiation	0.86					-
Teachers' Preferred Environment	120	Personalisation	0.78	-	0.52	0.07	0.40	0.42
		Participation	0.72		-	0.22	0.57	0.34
		Independence	0.82			-	0.33	0.09
		Investigation	0.86				-	0.36
		Differentiation	0.79					-
Teachers' Preferred Environment	81	Personalisation	0.80	-	0.55	−0.17	0.54	0.26
		Participation	0.78		-	0.08	0.58	0.23
		Independence	0.71			-	−0.22	−0.03
		Investigation	0.81				-	0.43
		Differentiation	0.62					-

both with the literature describing inquiry and open learning and with Moos' scheme for conceptualising human environments; it is economical in that it can be administered to a junior high school class in about 20 minutes; scales displayed satisfactory internal consistency and discriminant validity in three separate administrations; and the instrument has a large number of potential applications because of its four distinct formats (students' actual or preferred environment, and teachers' actual or preferred environment).

Currently the ICEQ is being employed in the investigation of the following research questions: the criterion validity of student perceptions in differentiating between classrooms following innovative and conventional curriculum materials; the incremental predictive validity of student perceptions in accounting for a substantial amount of learning outcome variance; the congruence of teachers' and students' perceptions of actual environment; the congruence between preferred and actual environment (either students' or teachers'): changes in teachers' perceptions of preferred learning environment during their first year of teaching; and determinants and effects of discrepancies between preferred and actual environment (either students' or teachers').

References

ADAMS, R.S. Perceived teaching styles. *Comparative Education Review*, 1970, 14, 50-59.

ANDERSON, G.J. *The assessment of learning environments: A manual for the Learning Environment Inventory and the My Class Inventory*. Halifax: Atlantic Institute of Education, 1973.

ANDERSON, G.J. & WALBERG, H.J. Learning environments. In H.J. Walberg (ed.), *Evaluating educational performance: A sourcebook of methods, instruments and examples.* Berkeley: McCutchan, 1974.

BARTH, R.S. Open education: Assumptions about learning and knowledge. Unpublished doctoral dissertation, Harvard Graduate School of Education, 1968.

BUSSIS, A.M. & CHITTENDEN, E.A. *Analysis of an approach to open education: Interim report.* Princeton: Educational Testing Service, 1970.

EBEL, R.L. Must all tests be valid? *American Psychologist*, 1961, 16, 640-647.

ELLIOTT, J. & ADELMAN, C. *The language and logic of informal teaching.* Ford Teaching Project, CARE, University of East Anglia, 1975.

FRASER, B.J. Classroom climate as predictor and criterion in science education research. *Research in Science Education*, 1976, 6, 109-120.

FRASER, B.J. Selection and validation of attitude scales for curriculum evaluation. *Science Education*, 1977, 61, 317-329.

FRASER, B.J. A curriculum evaluation. In H.J. Walberg (ed.), *Educational environments and effects: Evaluation and policy.* Chicago: McCutchon, 1978a. (in press).

FRASER, B.J. Measuring learning environment in individualized junior high school classrooms. *Science Education*, 1978b, 62, 125-133.

FRASER, B.J. & NORTHFIELD, J.R. A study of some aspects of ASEP during its first year of availability. Report to the Curriculum Development Centre, Canberra, 1976.

HOYLE, J.R. Are open space high schools more open? *Journal of Educational Research*, 1973, 67, 153-156.

HURD, P. DeH. *New directions in teaching secondary school science.* Chicago: Rand McNally, 1969.

INSEL, P.M. & MOOS, R.H. Psychological environments: Expanding the scope of human ecology. *American Psychologist*, 1974, 29, 174-188.

JOHNSON, D.W. & JOHNSON, R.T. *Learning together and alone: Cooperation, competition, and individualization.* Englewood Cliffs: Prentice-Hall, 1975.

LAZAROWITZ, R. & LEE, A.E. Measuring inquiry attitudes of secondary science teachers. *Journal of Research in Science Teaching*, 1976, 13, 455-460.

MARSH, C.J. A study of maintenance of some pedagogical attitudes held by social studies teachers after their first year of teaching in a school. *Australian Journal of Education*, 1976, 20, 316-324.

MASSIALAS, B.G. & COX, C.B. *Inquiry in social studies.* New York: McGraw-Hill, 1966.

MOOS, R.H. The assessment of the social climates of correctional institutions. *Journal of Research in Crime and Delinquency*, 1968, 5, 174-188.

MOOS, R.H. Conceptualizations of human environments. *American Psychologist*, 1973, 28, 652-655.

MOOS, R.H. *Evaluating treatment environments: A social ecological approach.* New York: Wiley, 1974.

MOOS, R.H. *Evaluating correctional and community settings.* New York: Wiley, 1975.

MOOS, R.H. & HOUTS, P.S. Assessment of the social atmospheres of psychiatric wards. *Journal of Abnormal Psychology*, 1968, 73, 595-604.

MOOS, R.H. & INSEL, P.M. (eds). *Issues in social ecology: Human milieus.* Palo Alto: National Press Books, 1974.

MOOS, R.H. & MOOS, B. A typology of family social environments. *Family Process*, 1976, 15, 357-372.

MOOS, R.H. & TRICKETT, E.J. *Classroom environment scale manual*. Palo Alto: Consulting Psychologists Press, 1974.

NORTHFIELD, J.R. The effect of varying the mode of presentation of ASEP materials on pupils' perceptions of the classroom. *Research in Science Education*, 1976, 6, 63–71.

POWER, C.N. & TISHER, R.P. Variations in the environment of self-paced science classrooms: Their nature, determinants, and effects. Unpublished paper presented to Australian Association for Research in Education, Adelaide, 1975.

POWER, C.N. & TISHER, R.P. Relationships between classroom behavior and instructional outcomes in an individualized science program. *Journal of Research in Science Teaching*, 1976, 13, 489–497.

SILBERGELD, S., KOENIG, G.R. & MANDERSCHEID, R.W. Assessment of the psychosocial environment of the classroom: The Class Atmosphere Scale. *Journal of Social Psychology*, 1976, 100, 65–76.

STEELE, J., HOUSE, E. & KERINS, T. An instrument for assessing instructional climate through low-inference student judgements. *American Educational Research Journal*, 1971, 7, 447–466.

TISHER, R.P. & POWER, C.N. The effects of classroom activities, pupils' perceptions and educational values in lessons where self-paced curricula are used. Report to the Australian Advisory Committee on Research and Development in Education, 1975.

TISHER, R.P. & POWER, C.N. Variations between ASEP and conventional learning environments. *Australian Science Teachers Journal*, 1976, 21(1), 27–29.

TRAUB, R.E., WEISS, J., FISHER, C.W. & MUSELLA, D. Closure on openness: Describing and quantifying open education. *Interchange*, 1972, 3, 69–84.

TRICKETT, E.J. & MOOS, R.H. Social environment of junior high and high school classrooms. *Journal of Educational Psychology*, 1973, 65, 93–102.

TRICKETT, E.J. & MOOS, R.H. Personal correlates of contrasting environments: Student satisfaction in high school classrooms. *American Journal of Community Psychology*, 1974, 2, 1–13.

WALBERG, H.J. The psychology of learning environments: Behavioral, structural, or perceptual? In L.S. Shulman (ed.), *Review of research in education*, (Vol. 4), Itasca, Ill.: Peacock, 1976.

WALBERG, H.J. *Educational environments and effects: Evaluation and policy*. Chicago: McCutchan, 1978 (in press).

WALBERG, H.J., SINGH, R. & RASHER, S.P. Predictive validity of student perceptions: A cross-cultural replication. *American Educational Research Journal*, 1977, 14, 45–49.

WALBERG, H.J. & THOMAS, S.C. Open education: An operational definition and validation in Great Britain and the United States. *American Educational Research Journal*, 1972, 9, 197–208.

Research in Science Education, 1978, 8, 90-98

THE INTERACTION OF STUDENT CHARACTERISTICS AND SCIENCE TEACHING ON STUDENT PERCEPTION OF THE LEARNING ENVIRONMENT

Adrian Fordham

Introduction

This paper is concerned with student perceptions of teacher behaviour and the learning environment, and how such perceptions are influenced by particular characteristics of the student.

The focus throughout is on classroom interaction, and emphasises the view that student behaviour, whether it be perceiving, remembering or problem solving, is an outcome of the relationship between the student and his learning environment.

The needs-press model of Murray (1938) is a good example of this interactive perspective and one that has been related to Australian education practice in general (e.g. Choo, 1972) and science education in particular (e.g. Gardner, 1972; Fordham, 1977). In this model a person's behaviour may be considered as a function of both internalised needs and the pressive characteristics of the surrounding environment with which the person is interacting. This paper considers one such needs-press complex, viz. achievement-orientated behaviour. There are three basic properties of this complex. Firstly there is the predispositional property of the student to act in a particular achievement orientated manner. Secondly there are varying states of arousal which will be dependent upon the achievement press of the environment. This press will consist of those gesticulations, comments, facial expressions, behavioural expectations etc. that constitute the student's social (school and classroom) environment and which have particular relevance to the arousal of achievement motivation. Thirdly there are *goals* to which achievement orientated activity is directed. These goals are not sets of discrete problems, definitions etc. but rather abstractions from those activities presented by the teacher and with which the student interacts. Two such goals have been proposed (Fordham, 1977). The first is an intergrated knowledge of the curriculum where the constituent topics and concepts are meaningfully related to each other and to broader areas of interest. The second is a non-integrative type of knowledge where the constituent elements remain independent of each other and are neither generalised nor related to other forms of knowledge. The teacher's specification of the former is referred to as 'emphasis on integration' and the latter as 'emphasis on specific detail'.

What then might be the effect of achievement motivation upon the perception of the 'pressive' characteristics of the environment, as well as the learning goals specified by the teacher? One might expect that if the environment is indeed 'pressive' then those students of higher achievement motivation who 'frequently seek for [that] certain press' (Murray, 1938), would construct a different, and more pressive, knowledge of that environment than students of lower achievement motivation. Further, one might expect a similar interactive relationship to exist in regard to the goals to which achievement behaviour is directed.

A second interactive relationship relevant to classroom practices is that between the

---

This paper forms part of a larger study done by the author while a member of the Education Research Unit, Research School of Social Sciences, Australian National University.

student characteristic 'intrinsic motivation' and those aspects of the learning environment most likely to lead to its arousal.

Intrinsic motivation can be defined as the tendency of an individual, in this case a student, to create, maintain and then resolve cognitive conflict. (See for example the papers of the OISE Symposium on Intrinsic Motivation, Day, Berlyne & Hunt, 1971). However the student will not exhibit this cognitive strategy unless in an aroused state. Arousal will be facilitated by the responsiveness of the classroom environment to the intellectual needs of the student. This includes both the provision of a diversity of information that is of interest to the student and likely to result in the generation of cognitive conflict as well as sufficient intellectual freedom to resolve such conflicts in an integrative manner. This dimension of the learning environment has been referred to as 'facilitation of intrinsic motivation' (Fordham, 1977).

Just as the achievement orientated student may frequently seek out pressive conditions, so too the intrinsic motivated student is more likely to seek out novelty and be in situations where conflict arises (Beswick, 1974). Hence one would expect such behaviour to influence his knowledge of the learning environment which he constructs. In particular one would expect that students of higher intrinsic motivation, when placed in a learning environment likely to facilitate the arousal of intrinsic motivation, to be more discriminating in regard to both novel experiences and the opportunity for intellectual freedom than students of lower intrinsic motivation.

A third student characteristic that is relevant to the present paper is the student's level of 'cognitive readiness'. This refers to the student's prior development of those cognitive structures necessary for learning a particular section of the curriculum. Students vary in their possession of these cognitive structures and hence in their readiness to assimilate or incorporate novel aspects of the curriculum. Since the student's construction of a knowledge of that curriculum is dependent, to a large extent, on his level of cognitive readiness, one would expect students of differing levels of cognitive readiness to perceive the teacher's specification of each of the two learning goals described above quite differently. This would be dependent on the extent to which the students find each of these learning goals satisfying and challenging.

In the results presented below only those analyses have been included that bear directly upon the relationships between the student and his learning environment that have been discussed above.

Sample

The sample of students and teachers used for the analyses described in this paper comprised those 5th form biology classes in the Australian Capital Territory that were due to commence Chapter 6 (Living in Water) and Chapter 7 (Living on Land) of the *Biological Science*: *Web of Life* Curriculum at, or shortly after, the beginning of second term, 1975. As a result the final sample comprised seventeen classes numbering 274 students from whom complete data sets had been collected.

The Measurement of the Student Characteristics

Cognitive readiness

A list of thirty-two prerequisite concepts that are necessary for the development of an understanding of 'Living in Water' and 'Living on Land' was identified. A Cognitive Readiness Test was constructed to ascertain the students' prior acquisition of these concepts and administered immediately prior to the commencement of 'Living in Water'. The thirty item test possessed a K-R20 reliability coefficient of 0.82.

Need for achievement

The achievement scale contained within the Activities Index (Stern, 1970) was used, particularly because its use by Gardner (1972) and Choo (1972) has demonstrated its suitability within the context of the Australian secondary school system. The reliability coefficient of the 10-item scale was 0.76.

Intrinsic motivation

The short Intrinsic Motivation (Curiosity) Scale consisting of sixteen items constructed by Beswick (1974) was used. This short intrinsic motivation scale has been used at both the tertiary and secondary levels of education in Australia and has reported reliabilities of 0.71 and 0.68 respectively (Beswick, 1974, 1975). In the present study its reliability coefficient was 0.65.

The Assessment of the Learning Environment

Two assessments were made of the learning environment, the first by the researcher and the second by individual students.

The student assessment of the learning environment

A questionnaire procedure was used to assess individual student's perceptions of the learning environment along each of the relevant dimensions. In the case of achievement press, the researcher was of the opinion that the overall school environment as a source of arousal of achievement motivation was more important than a specific classroom. Stern's Achievement Press scale (Stern, 1970) was used to assess this dimension. Three scales were constructed, based on the Biology Classrooms Activities Checklist (Best, 1974), to assess the remaining three dimensions of the learning environment: emphasis on integration (10 items, reliability coefficient 0.76), emphasis on specific detail (6 items, reliability coefficient 0.62), facilitation of intrinsic motivation (10 items, reliability coefficient 0.76). Each of these scales are referred to as classroom perception scales.

The researcher's assessment of the learning environment

The methods used for this assessment of the learning environment are detailed in Fordham (1977) and are only briefly commented upon here. The procedure used for the assessment of the dimension facilitation of intrinsic motivation, emphasis on specific detail and emphasis on integration utilised structured student interviews, informal discussions with teachers and students and class observation work. An overall assessment of the learning environment was made following the conclusion of 'Living on Land'. It consisted of the researcher taking into account the total set of observations, interviews and informal discussions that resulted from this procedure and making a relative judgment on the basis of this experiential data. In the case of the specification of learning goals, this procedure resulted in the teachers being rated on a 1-7 scale; with the facilitation of intrinsic motivation a 1-6 scale was appropriate. This degree of discrimination was the maximum justified by the procedure.

Since the achievement press of the school was assessed, rather than that of the classroom, the researcher decided to use the class mean scores on the Stern achievement press scale (see above) for the purpose of the analyses reported in this paper.

Findings

Two models were used to examine the six relationships described above between the student characteristics and student perceptions of the learning environment.

(i) the student's perception of his learning environment could be considered an additive function of both the personality variables that characterise the student and the environment itself.

This understanding of the relationship closely resembles the notion of 'projection' described by Stern (1970) and Gardner (1975). However the interactive perspective just outlined leads to a different view of the possible relationship between student characteristics and student perceptions of the learning environment. It suggests that:

(ii) the student's perception of his learning environment could be considered an interactive function of both the personality variables that characterise the student and the environment itself.

Student perceptions of their learning environment – an additive model

A multiple stepwise regression analysis was used to examine the proposed relationships between the students' levels of intrinsic motivation, achievement motivation, and cognitive readiness and their perceptions of the relevant dimensions of the learning environment. The following 2-term regression model was proposed:

$$\begin{array}{l} \text{Student score on} \\ \text{particular classroom} \\ \text{perception scale} \end{array} = \beta_1 \begin{array}{l} \text{researcher's}^{1} \\ \text{rating of} \\ \text{corresponding} \\ \text{classroom} \\ \text{dimension} \end{array} + \beta_2 \begin{array}{l} \text{student} \\ \text{characteristic} \end{array}$$

The simple correlations between the students' perceptions of their learning environment and each of the student characteristics in the six proposed relationships are reported in Table 1.

---

[1] In the case of achievement press this was the consensual (class mean) press score.

TABLE 1

Correlations between the Students' Levels of Intrinsic Motivation, Achievement Motivation, Cognitive Readiness, and their Perceptions of Four Dimensions of the Learning Environment

Student Characteristics	*Classroom Perception Scale* *Emphasis on Integration*	*Emphasis on Specific Detail*	*Facilitation of Intrinsic Motivation*	*Achievement Press*
Intrinsic Motivation			0.09	
Achievement Motivation	0.14^{*}	0.01		0.05
Cognitive Readiness	−0.05	-0.24^{*}		

* $p < 0.05$

After controlling for the effects of the researcher's rating of the relevant dimensions of the classroom environment, only two significant relationships were present, of which one was quite weak. The students' perception of their teacher's emphasis on specific detail was negatively related to the students' level of cognitive readiness: students of low cognitive readiness tended to perceive their teacher's presentation of the curriculum as placing greater emphasis on the retention of specific detail than did their peers of higher cognitive readiness (partial correlation = 0.28). There was a weak, positive relationship between the students' perception of their teacher's emphasis on integration and the students' level of achievement motivation (partial correlation = 0.13).

Student perceptions of their learning environment – the interactive model

Two procedures were used to investigate the relationships between student characteristics and student perceptions of the learning environment in terms of an interactive model.

The first was a stepwise multiple regression procedure, similar to that above, but with a 3-term regression model including the interactive term 'researcher's rating * student characteristic'. Because of multicollinearity, this procedure was appropriate for an examination of only three of the proposed relationships, viz. the relationships between the student's level of cognitive readiness and his perception of each of the learning goals; and finally the relationship between the student's level of achievement motivation and emphasis on integration.

This multiple regression procedure demonstrated that the student's perception of his teacher's emphasis on integration was a function of the interaction of his level of cognitive readiness and the researcher's rating of the environment on the corresponding dimension ($6.31 > 3.86 = F(1,272)0.05$). This relationship is represented in Figure 1. Consider the differences in student perception of those learning environments rated by the researcher high on the dimension emphasis on integration. This corresponds to slope 'a' in Figure 1. The positive gradient of slope 'a' indicates that students of higher cognitive readiness perceived these classes

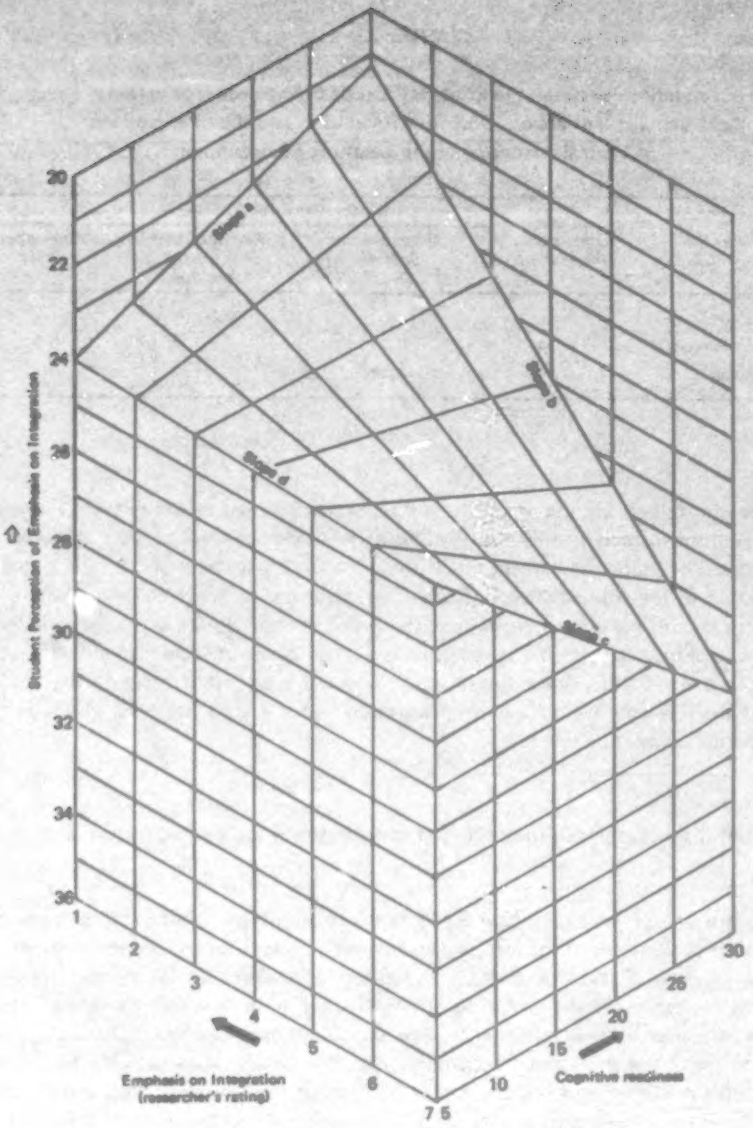


FIGURE 1

The regression surface representing the interaction of cognitive readiness and the researcher's rating of the dimension emphasis on integration upon the students' perception of that dimension. The regression equation, using raw regression weights, was as follows:

Student perception = $24.8 - 0.23x - 0.17y + 0.06xy$ where

x = researcher's rating

y = cognitive readiness

as characterised by a greater emphasis on integration than did students of lower cognitive readiness. Contrast this with those classes rated 'low' on this dimension by the researcher which corresponds to slope 'c'. Students of higher cognitive readiness perceived these classes as being far lower in their emphasis on integration than did students of lower cognitive readiness. Finally a comparison of slopes 'b' and 'd' indicates that students of higher cognitive readiness were far better able to discriminate between classes characterised by different emphases on integration than were students of lower cognitive readiness.

There was no significant interactive effects in the case of each of the remaining relationships tested using this procedure.

For those three relationships where the above procedure was inappropriate the following form of analysis was adopted. Firstly the classes were divided at the median on each dimension into two categories, 'high' or 'low'. Secondly the correlations between the students' scores on the classroom perception scales and their scores on each of the student characteristic measures were calculated. This was done for students found in classes grouped 'high' on the corresponding classroom dimensions and then for those students found in classes grouped 'low'. Thirdly a multiple regression procedure with dummy variables was used to test the heterogeneity of each pair of correlation coefficients.

Each of the three relationships examined using this procedure was found to be significant. Firstly for students found in classes characterised by a high level of facilitation of intrinsic motivation, those students of higher intrinsic motivation perceived classes as higher on this dimension than did students of lower intrinsic motivation; no such relationship existed for those students found in classes characterised by a low level of facilitation of intrinsic motivation ($r = 0.22, -0.04; 4.89 > 3.86 = F(1,272)0.05$). A similar relationship existed for the students' perception of the achievement press of the school. Students of higher achievement motivation rated their learning environments higher on the dimension achievement press than students of lower achievement motivation, but only in those classes where the consensus was that the achievement press was in fact high ($r = 0.15, -0.08; 3.98 > 3.86 = F(1,272)0.05$). Further students of higher achievement motivation rated their learning environment higher on the dimension emphasis on specific detail than did students of lower achievement motivation only in those classes grouped as 'high' emphasis on specific detail ($r = 0.16, -0.09; 4.68 > 3.86 = F(1,272)0.05$).

Discussion of Results

The analyses supported the proposition that students under certain conditions tended to construct a knowledge of the 'pressive' characteristics of their learning environment that was dependent on the students' levels of achievement motivation. Stern (1970) has described such a finding as a 'projection effect' and argued, as does Gardner (1975), that if such a process occurred it would raise doubts about the validity of students reporting environmental characteristics. Both Stern's and Gardner's objection is methodological and based upon their attempts to measure 'objectively' the students' learning environments. However in this study such an effect was predicted – not as a main effect but rather an interactive one.

Similarly the analyses support the interactive relationship proposed earlier, between the student's level of intrinsic motivation and his perception of the learning environment's facilitation of intrinsic motivated behaviour.

The student interview data assisted in the interpretation of the interactive relationship between level of achievement motivation and teacher emphasis on specific detail. Academic performance at the HSC examination was thought by many students to be related to the teacher's emphasis on specific detail. Students in classes rated low on the dimension teacher emphasis on specific detail often commented that their teacher should spend more time marking guide questions, test their knowledge of the textbook regularly, and only deal with material and activities directly concerned with the prescribed curriculum and especially the textbook. By so doing, these students felt that their chances of success at the HSC examination would be increased.

Evidently these students had defined the form of biological knowledge to be evaluated at the HSC examination as requiring the retention of those specific details dealt with in the *'Web of Life'* curriculum. One might expect then that the specification of such a learning goal would possess 'pressive' characteristics similar to the achievement press of the school and be arousing for students of higher achievement motivation.

The qualitative data did not assist in the interpretation of the weak main effect of achievement motivation upon 'emphasis on integration'. Since such a finding was unexpected, little credence can be placed upon it without further investigation.

The effect of cognitive readiness upon the students' perceptions of each of the learning goals parallels the work of Tamir on the relation between achievement and cognitive preference (Tamir, 1976).

Students of higher cognitive readiness found in classes with a strong emphasis on the retention of specific detail often expressed boredom with the curriculum as it was presented – the extent of this boredom is apparent in the following student's description of her learning experience:

> . . . writing out all the questions and answers, and all that again, going into all the detail, you might know it but you have to write it out again to be handed in and that takes up all the time . . . it's the dull routine . . . it's knowing what we are to be doing day after day after day.

In contrast, those students of low cognitive readiness expressed satisfaction with this learning goal. When found in classes where there was no such emphasis, most students of low cognitive readiness commented that there was insufficient direction given by the teacher; that insufficient notes were dictated to them; and that there was too little emphasis placed upon the marking of guide questions.

A learning environment characterised by a high emphasis on the retention of specific detail is an environment in which students of low cognitive readiness can intellectually cope, and the attainment of such a goal specified by the teacher can be considered satisfying to these students. This then becomes their preferred mode of learning or information processing. Hence those cues that relate to the retention of specific details will be particularly salient to these students, and they will tend to perceive their learning environment stronger on this dimension than do students of higher cognitive readiness.

The interview data revealed a different pattern of results when students commented upon the level of teacher emphasis on integration that prevailed in their classroom. For those classes where there was a high level of teacher emphasis upon integration two types of responses were represented. Again these were generally dependent upon the students' level of cognitive readiness. Most students of lower cognitive readiness stated that they could not see the relevance of much of the material dealt with in the class, and generally felt unable to cope with the learning goal specified by the teacher. Rather they tended to see this learning goal more suited to the 'better' students. Indeed the student of higher cognitive readiness did tend to find this environment stimulating, often commenting that it was better than the 'traditional' methods. This difference in attitude, by students of differing cognitive readiness, towards the teacher's presentation of this learning goal is best expressed by the following student of low cognitive readiness:

> ... the practs ... you have to be hypothetical and I suppose if you've got brains this is pretty challenging and pretty good fun but if you're not really top of the class, a struggler, it's a bit hard trying to think out how things could happen ...

For students of higher cognitive readiness, relationships inherent in the specification of an integrative learning goal are both meaningful and attainable. The teacher's specification of this learning goal is congruent with their preferred mode of learning; yet the quantitative data indicate that this does not result in a main effect of the sort just described. In particular the analyses indicated that students of higher cognitive readiness were better able to discriminate between classes of differing levels of emphasis on integration than were students of lower cognitive readiness. This is to be expected. Students of lower cognitive readiness, lacking necessary prerequisite concepts, are unable to construct a knowledge of those relationships characteristic of an integrative learning goal. To put it another way, they are unable to construct a basis for comparison. In contrast, students of higher cognitive readiness do possess the necessary prerequisite concepts to speculate upon what such potential relationships might be; as a result they are able to more accurately rate their teacher's presentation of the cirriculum on this dimension.

Conclusion

The interactive perspective adopted in this study suggested an alternative approach to the study of the effects of student characteristics upon their perceptions of the classroom. This perspective emphasised the necessity to examine the nature of the interaction between the student (or more exactly the student variable) and his learning environment before testing for the presence of these effects. Six relationships were consequently proposed, and both the analyses and qualitative data supported the presence of five of these in the ongoing activities of the classroom.

References

BEST, E.D. Student perceptions of activities and outcomes in South Australian matriculation biology classes. *Research in Science Education,* 1974, 4, 151-160.

BESWICK, D.G. Intrinsic motivation in senior secondary school students. *Education Research and Perspectives,* 1974, 1 (2), 15-25.

BESWICK, D.G. The students' perspective: Introduction to the surveys. In D.S. Anderson, K.J. Batt, D.G. Beswick, G.S. Harman & C. Selby Smith (eds), *Regional colleges: A study of non metropolitan colleges of advanced education in Australia.* Australian National University: Education Research Unit (3 vols).

CHOO, P.F. The Need-Press Model: Its application to academic performance and aspirations among high school students. Unpublished Ph.D. thesis, University of Western Australia, 1972.

DAY, H.I., BERLYNE, D.E. & HUNT, D.E. (eds), *Intrinsic motivation.* Canada: Holt, Rinehart & Winston, 1971.

FORDHAM, A.M. The interaction of student motivation and teacher behaviour upon academic performance. Unpublished Ph.D. thesis, Australian National University, 1977.

GARDNER, P.L. Attitudes to physics. Unpublished Ph.D. thesis, Monash University, 1972.

GARDNER, P.L. Projection effects and the Needs-Press Model. *The Journal of Social Psychology,* 1975, 95, 91-97.

MURRAY, H.A. *Exploration in personality.* New York: Oxford University Press, 1938.

STERN, G.G. *People in context.* New York: John Wiley, 1970.

TAMIR, P. The relationship between achievement in biology and cognitive preference styles in high school students. *British Journal of Educational Psychology,* 1974, 46, 57-67.

Research in Science Education, 1978, 8, 99-109

SCIENCE FACILITIES AND VARIETY IN SCIENCE TEACHING

John Ainley

Introduction

A premise on which many government programs of assistance to schools are based is that improving physical facilities will lead to an improvement in the quality of education provided. However, despite the fact that governments proceed as if this premise were true there is little systematic evidence to support it. Indeed among teachers, parents and citizens contrary views are expressed about the importance of good facilities. The view accepted in this paper is that expressed by Dent:

> First and foremost and all the time in any education system, it is the teachers who matter most. But it is the gravest error to say as is often said: Teachers matter but not the buildings. In fact, the buildings matter a very great deal. Good buildings, well equipped, make both teachers and pupils work easier as well as more agreeable.
>
> (Dent, 1971, p. 139)

In this paper the relationship between the standard of science facilities and student reports of the stimulation and variety in their learning environment is explored. The study arises from a larger evaluation of the Australian Science Facilities Program (Ainley, 1977a). Under that Program some $123 million was provided for the improvement of science facilities in secondary schools. Those who were responsible for the program intended that, in addition to improving the material facilities for science teaching, the Program would lead to an improvement in the quality of science education and possibly greater student understanding of, and interest in science.

Such expectations can be considered in terms of the general curriculum model elaborated by Dahllöf (1971). In that model achievement was considered in relation to objectives, frame factors and processes. The specific instance considered in this paper is the relationship between one frame factor – facilities – and one aspect of classroom processes – stimulation through variety.

The importance of facilities: General issues

Environmental Psychology

General interest in the relationship between the physical environment and human behaviour has been evident in several contexts. Proshansky, Ittleson and Rivlin (1970) have collected reports of a number of studies concerned with this area of research. Their acknowledgment that the field lacks a theoretical structure is in agreement with the conclusion reached by Drew (1971) as a result of a major review. However, there is a general awareness by these writers that changes in the physical environment might produce changes in behaviour even though the mechanisms are not understood. Other publications have suggested important methodological issues regarding research in this field. Barker (1968) argued that the aggregate of participants, rather than individuals, should be the unit of analysis in studies in this area. Proshansky (1974) contended that a major task in environmental psychology was to establish relationships between 'the setting as it is' and 'the setting as it is experienced'. Both Barker (1968) and Proshansky (1974) argue for systematic objective approaches to the description of the physical setting, claiming that this destination has not always been clear.

Studies of teaching methods in science

Facilities have been reported in some studies as acting as constraints on certain teaching methods (e.g. Holley, 1974). This view of the role of facilities was broadened by Englehardt (1968) who argued that in addition to acting as constraints facilities could influence teaching methods in other ways. He proposed the term 'suggestive space' to embody the notion that some methods of teaching might be stimulated by certain spatial characteristics of science rooms. After a study of the methods used by 500 teachers in 59 schools in the U.S.A., Englehardt (1968) identified some characteristics of science rooms that appeared to be associated with the use of enquiry methods in science. The type of science room was important; and especially important when teachers had adequate preparation time, easy access to the preparation area and taught mainly in one room each. There were two important methodological features of Englehardt's study. The first was that the examination of spatial influence included consideration of necessary features as well as design differences. The second was that it was conducted in a specific context so that it was possible to choose dependent variables most likely to be affected by the facilities available.

Studies of teaching methods in areas other than science

In general the results obtained from research on the influence of general school facilities have been less conclusive than those from research on specialist science facilities. One recent study by Weinstein (1977) suggested that the behaviour of primary school students could be modified through changes in the physical design of an open classroom. In this study the spatial distribution of activity was observed before and after a change in physical design. The changes made were specific and related to the behavioural goals which were tested. For example changes in the science area increased the amount of use pupils made of that area. A general point which is suggested by this study is that the influence of the physical environment on student activity is most likely to be detected when specific aspects of design are studied in relation to logically connected aspects of behaviour.

Stimulation through variety

The concept

The concerns of the Commonwealth Standards Committee regarding the quality of science teaching were discussed in a previous article (Ainley, 1976). In brief the members of that committee appeared to be concerned with both the environment in which science was studied and the nature of the activities which took place in science lessons. Three concepts seemed to be important components of the learning environment: individual involvement, organization, and stimulation through variety. Similarly there were three important aspects of the activities in which students engaged, which the Committee believed could be nurtured by better facilities. These were the encouragement given to students to explore a range of scientific phenomena, the emphasis on practical work, and a more limited use of textual materials. In the present paper attention has been focussed on one aspect of the learning environment: stimulation through variety. Stimulation was seen as coming through both the provision for varied and flexible instruction and provision of a wide variety of stimulating materials in the room. In its final report the Committee stated that a science room should be suitable for all kinds of activity including *inter alia*, demonstrations, pupil experiments, group discussions, films, or a mixture of activities (Commonwealth of Australia, 1976, p. 16).

This aspect of the learning environment had been used in other studies of teaching behaviour and classroom processes. One of three dimensions of teacher behaviour identified by Ryans (1960) in a factor analytic study of some 6000 teachers was called 'Pattern Z'. 'Pattern Z' was defined as 'stimulating, imaginative, surgent, versus dull, routine teacher behaviour' (Ryans, 1960, p. 382). This was a similar dimension to that which Anthony (1967) called Stimulation. Anthony's measure of stimulation included assessing variety in materials, techniques and feedback to students. Keeves (1971) also used a variable called 'stimulation for learning' which incorporated measures of the variety of methods used.

Similar concepts to this were assessed by two sets of scales designed to assess student perceptions of the learning environment. A Classroom Environment Scale developed by Trickett and Moos (1974) included a subscale designated Innovation. This was described as measuring the extent to which different modes of teaching and interaction occurred in the class (Trickett & Moos, 1974, p. 181). A more appropriate title would appear to be variety: a title which makes the similarity to the concept of stimulation through variety more obvious.

The Learning Environment Inventory (Anderson & Walberg, 1974) is a well known instrument for assessing learning environments, using the opinions of students. That inventory contained scales which were a little different from those used by Trickett and Moos (1974), and which were not so clearly related to the concepts at present being considered. Nevertheless the scale called 'diversity' contained elements which were rather similar to the concept of stimulation through variety.

The proposition

As a result of considering the assumptions of the Australian Science Facilities Program and general literature concerned with the provision of educational facilities it was possible to formulate a proposition to guide the investigation.

> **Student reports of greater stimulation through variety will be related to better standards of science facilities.**

The measure

Evidence about the stimulation through variety was obtained from students' answers to a seven item scale which formed part of a larger questionnaire.

This and other scales on the questionnaire have been described elsewhere (Ainley, 1977a). It was decided that a questionnaire would be a better means of obtaining information about the aspect of science lessons than direct observation. The main reason for this decision was that it was the student perceptions of their science lessons which were of greatest interest. In addition there were three other considerations.

1. In answering questions students could consider a larger sample of lessons than would be possible for an observer.

2. Students could consider all the activities in the class and not only those directly visible or those involving the teacher.
3. Students would be able to take into account subtle nuances in teaching procedures to which an objective observer would be less sensitive.

Nearly 3000 students from 105 classes completed the questionnaire which was administered during visits to 29 schools. Schools were chosen so that a full range of standards of science facilities was represented. The questionnaire was designed for, and administered to Year 9 students because it was believed that the effects of facilities on teaching would be greatest at that level.

The unit of analysis

The class was considered to be the logical unit of analysis in this study. It was the conceptual unit of interest. As Wiley (1971, p. 264) has observed, when a teaching program is the object of an evaluation the class will usually be the appropriate vehicle for the evaluation. Barker (1968) has also stressed that in studies of the influence of the physical environment on behaviour, aggregates provide more appropriate units of analyses than individuals.

The effects of facilities on certain aspects of science lessons were the main concern of the present study. Each of these aspects of science lessons concerned the whole class. Since each class experienced its science lessons as a group it seemed reasonable that their questionnaire responses should be analysed as a group.

TABLE 1

Properties of the Scale Measuring Stimulation Through Variety in Science Lessons

	Between students	*Between classes*
Number of Items	7	7
Reliability (alpha)	0.56	0.66
Mean	19.0	18.9
Standard deviation	4.6	2.6
Range	7 to 34	14 to 28
N	2991	105

Properties of the scale

The Reliability of the scale was calculated in two ways. First, the values of coefficient alpha were calculated using student responses. Second, the class means for each item were calculated for each class and the reliability of the scale using class means was calculated. As suggested by Kane and Brennan (1977) the reliability of the class mean was higher than that of the individuals though not as high as would be calculated by Shaycoft (1962) or Wiley (1971). Reliabilities and other relevant statistics are presented in Table 1. Even though the reliability of this scale was lower than that of other scales used in the evaluation (Ainley, 1977b) it did meet the criteria specified by Nunally (1967) for research instruments.

Properties of the class means

A number of important issues were considered in the choice of the class as the unit of analysis. One concerned the degree to which students in the same class were in agreement about the extent to which the learning environment was stimulating. It was argued that it should not be expected that members of each class would be unanimous in their response to questions about this aspect of their science lessons. However, it did seem important to assess the degree to which classes were empirically distinct. Two statistics were calculated. Both depended on a one-way analysis of variance procedure in which the total variance was partitioned into two parts: the variance of class means about the grand mean and the variance of student scores about the class means. The first statistic was the F ratio. It can be seen from Table 2 that the F ratio was significant at a probability level better than one in a thousand. This indicated that class membership was a significant factor associated with student responses. The second statistic was the ratio of the total sums of squared deviations associated with variations between classes. As shown in Table 2 some 27 per cent of the total variance was due to difference.

TABLE 2

Properties of the Class Mean Scores on Stimulation Through Variety in Science Lessons

Consensus within classes:	F ratio	12.7 (p < 0.001
	Eta squared	0.27
Distribution:	Kolmogorov-Smirnov Z	0.85 (p < 0.1)
Correlation coefficient	Teacher x class mean	0.39 (p <0.001)

Even though the main concern of this study was the students' perception of the learning environment it was of interest to compare the students' views with those of their teachers. It can be seen from Table 2 that the views of students were shared, to a moderate degree, by their teachers. The correlation coefficient was significantly greater than zero, though it was not large. However it was greater than has often been reported in such comparisons. A more exhaustive comparison of teachers and students views will be published as a separate article in the future.

For the class mean scores to be used as a dependent variable in a regression analysis it is necessary that they be normally distributed. A Kolmogorov-Smirnov test was performed, and, as can be seen from Table 2, the distribution of the class means did not differ significantly from a normal distribution.

The independent variables

On the basis of a previously reported factor analysis (Ainley, 1976), it had been decided that school science facilities could be considered in four dimensions. These were the availability of rooms, the quality of those rooms, the abundance of apparatus, and a number of laboratory assistants. In studying the effect of facilities on the views students held of their science learning, four variables corresponding to these dimensions were used.

1. The occupancy of science rooms was measured simply as the percentage of science lessons held in science rooms.

2. The quality of the science room most often used was assessed by the school science co-ordinator. The measure of 'quality' was a score obtained by summing the school science co-ordinator's assessment of the availability and accessibility of 14 key features. The scores on that scale correlated well with an observer's rating ($r = 0.83$). The scale had been previously shown (Ainley, 1978) to discriminate between rooms considered to be of different quality and to have a satisfactory internal consistency ($\alpha = 0.68$).
3. The abundance of apparatus in each school was represented as a composite score derived from the replies from science co-ordinators about the adequacy of major, minor and expendable apparatus.
4. The provision of laboratory assistants was represented by the number of laboratory assistants per 1000 students in the school.

The method of analysis

In the present study it was planned to investigate the way in which variables representing the facilities available for teaching science were related to student reports of stimulating through variety in science lessons. The independent variables in this study were "attribute' rather than 'active' variables; that is they were measured rather than manipulated variables. Multiple regression appeared to be the most suitable method of analysis in this situation where cases were not randomly assigned to treatment groups.

It has been seen in Table 2 that the class mean scores on the dependent variable were normally distributed as required for this form of analysis. In addition a sample size of 105 given four independent variables reduced the possibility that the regression weights would be unstable.

The regression analysis was conducted so that both main effects and interactions could be examined. In the first step of the analysis only the main effects were considered. The independent variables were entered in the regression analysis in a predetermined sequence: occupancy, quality, apparatus and assistants. The order in which the variables were entered corresponded to the priorities followed in the Science Facilities Program. In subsequent steps interaction effects were examined. In searching for interaction effects product variables were used. Since there were four independent variables there were six possible second order products, four possible third order products, four possible third order products, and one fourth order product. As recommended by Finn (1974, p. 95) all the predictor variables were standardized prior to the analysis as means of preventing scaling differences distorting the contribution of some variables.

In the first step of the analysis all four predictor variables were included in the order described above. Secondly, the second order product terms were entered in a stepwise procedure. The stepwise regression was stopped when an added term failed to reach significance at the five per cent level. Then the third order product terms were entered, and finally the fourth order interaction was considered after all other significant factors and combinations had been included.

Results

Main effects

The results of the regression analysis involving only main effects are shown in Table 3.

TABLE 3

The association of the standard of facilities with stimulation through variety: Results of the regression analysis

	Simple Correlation Coefficient r	*Standardized Regression Coefficient* β	*F for Regression Coefficient*	*Coefficient of Contribution* $c = r\beta$
Occupancy	*0.33*	0.11	0.93	0.04
Quality	*0.34*	0.18	2.60	0.06
Apparatus	*0.40*	*0.32*	10.29**	0.12
Assistants	−0.16	−0.01	0.67	0.00

Multiple $R = 0.48$ $R^2 = 0.23$

Overall $F = 7.31^{***}$ $SD_V = 2.51$

$\bar{V} = 19.07$

$N = 105$

Note: Coefficients significant at the five per cent level are in italics.
$^{**}\ p < 0.01$ $^{***}\ p < 0.001$

Simple correlation coefficients have been used to show the association between each predictor variables and the criterion. Standardized regression coefficients have been used to indicate the contribution of each variable to the criterion when the effects of other variables were controlled. An alternative way of presenting these data was to tabulate the coefficient of contribution for each variable. It is defined as the product of the simple correlation coefficient and the standardized regression coefficient. Even though its use has been disputed Keeves (1971, p. A200) has shown that it provides valuable information in cases where suppressor effects have not occurred. As might have been expected, the availability of apparatus was the most powerful predictor of classes reporting a stimulating environment in which a variety of methods and materials were used.

The results obtained demonstrate the value of using regression analysis for such a study as some variables for which there was a significant simple correlation with the criterion failed to reach significance in their regression coefficients. The standardized regression coefficients for the variable representing the availability of apparatus was the only regression coefficient significant at the five per cent level or better. While room quality and room occupany had significant correlations with stimulation, that association diminished when the effects of other variables were controlled. It should be noted that room quality, and, to a lesser extent, the frequency of science lessons being in science rooms, also contributed to the explained variance on the scale measuring the degree of stimulation in the learning environment. The coefficient of contribution indicated that availability of apparatus accounted for some 12 per cent of the variance in this scale. From the value of the multiple correlation coefficient it can be seen that these measures of the standard of facilities in science rooms account for 23 per cent of the scale variance.

Interaction effects

There were three significant second order interaction terms from the regression analysis in which stimulation was the criterion. These interactions were between apparatus availability and the provision of assistants, apparatus availability and room quality, and room occupancy and room quality. The results of the regression analysis have been presented in Table 4.

TABLE 4

The association of the standard of facilities with stimulation through variety: Regression analysis with interaction

Block	*Variable*	*Simple Correlation Coefficient*	*Standardized Regression Coefficient*	*F for Regression Term*
1	Occupancy	*0.33*	-0.06	0.19
	Quality	*0.34*	*0.32*	8.40**
	Apparatus	*0.40*	*0.36*	15.63***
	Assistants	-0.16	*0.26*	6.47*
	Apparatus x Assts.	*0.28*	*0.39*	17.20***
	Apparatus x Quality	0.04	*0.29*	11.16***
	Occupancy x Quality	*-0.31*	*-0.24*	3.99*

Multiple $R = 0.62$ $R^2 = 0.39$

Overall $F = 8.78^{***}$

Note: Coefficients significant at the five per cent level are in italics.
* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Inclusion of the interaction terms increased the multiple correlation considerably. The percentage of the total variance in stimulation attributable to the standard of facilities rose to 39 per cent: an increase of 16 per cent over what could be explained by the main effects alone.

The magnitude of the regression coefficients were modified by the inclusion of interaction terms as can be seen in the results reported in Table 4. It can be seen that the main effect term for apparatus remained important but the regression coefficients for room quality and assistants as independent terms increased somewhat.

The three effects of the interaction terms are represented in Figures 1, 2 and 3. It can be seen from Figure 1 that increasing the abundance of apparatus resulted in a more stimulating environment only when there was ample provision of laboratory assistants. Figure 2 indicates that an abundance of apparatus had greater impact on the degree of stimulation in the environment of good quality rooms than in poor quality rooms. It would appear that more apparatus promoted the stimulation provided in better quality rooms to a greater extent than it compensated for the deficiencies in poor rooms.

The interaction between room occupancy and room quality with respect to stimulation through variety is presented in Figure 3. When all science lessons were in sciene rooms there was little difference between good and poor quality rooms. However when some lessons were not in science rooms the degree of stimulation reported by those students in good quality rooms was greater than their peers in poor quality rooms. One plausible explanation for this effect would be in the nature of the activities mentioned in the stimulation scale. A teacher who made use of the varied resources and approaches to teaching which were mentioned

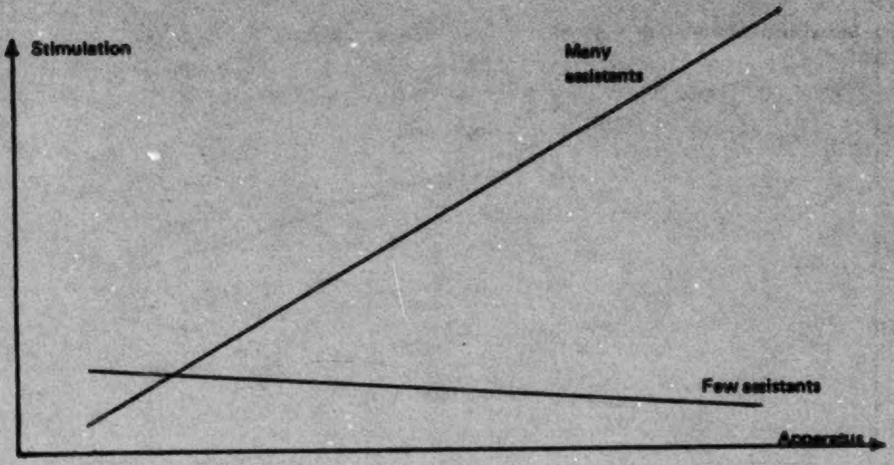


FIGURE 1

The interaction of apparatus with assistants on stimulation

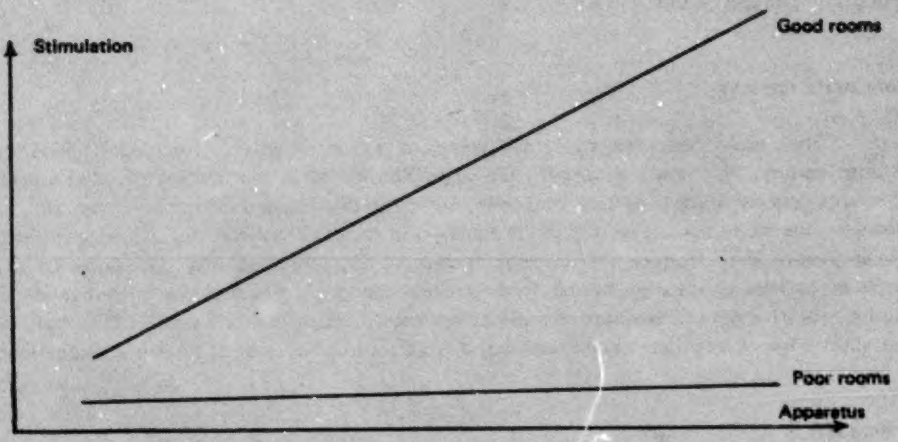


FIGURE 2

The interaction of apparatus with room quality on stimulation

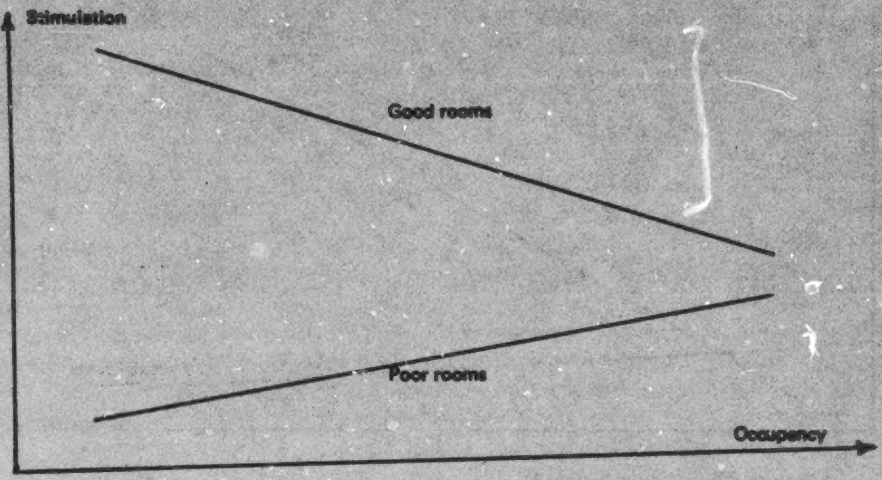


FIGURE 3

The interaction of occupancy with room quality on stimulation

might also make use of resources out of science rooms. Hence, it could be a deliberate policy to conduct some lessons out of science rooms.

The study being reported not only used data collected by means of questionnaires answered by students, but also involved visits to schools to form a set of case studies. This case study data can be found in Ainley (1977a).

Concluding remarks

This paper has been concerned with one aspect of science teaching: stimulation through variety. The results reported need to be considered in the context of other aspects of science lessons which have been studied in relation to the standard facilities (Ainley, 1977a). However, the study does suggest that the standard of facilities available did influence students perceptions of their learning environment in science. This influence may have occurred as a result of barriers to activities having been removed, changes in the attitudes or motivation of teachers, or changes in the responsiveness of students. Though the data used in this study do not allow these distinctions to be examined, they ought to be pursued in further investigations.

References

AINLEY, J.G. The Australian science facilities program: A study of its impact on schools. *Bulletin of the Victorian Institute of Educational Research*, 1976, 36, 17–39.

AINLEY, J.G. *The Australian science facilities program: A study of its influence on science education in Australian schools.* Melbourne: Australian Council for Educational Research, 1977a.

AINLEY, J.G. The impact of science facilities on science learning. *Research in Science Education*, 1977b, 7, 213-222.

ANDERSON, G.F. & WALBERG, H.J. Assessing classroom learning environments. In I. Marjoribanks (ed.), *Environments for learning.* London: National Foundation for Educational Research, 1974.

ANTHONY, B.C.M. *The identification and measurement of classroom environmental process variables related to academic achievement.* Unpublished doctoral dissertation, University of Chicago, 1967.

BARKER, R. *Environmental psychology.* Stanford, Calif.: Stanford University Press, 1968.

COMMONWEALTH OF AUSTRALIA. *Commonwealth advisory committee on standards for science facilities in independent secondary schools* (Chairman, E.D. Gardiner). *Final Report, 1976.* Canberra: Department of Education, 1976.

DAHLLÖF, U. Relevance and fitness analysis in comparative education. *Scandinavian Journal of Educational Research*, 1971, 15(3), 101-121.

DENT, H.C. *1870-1970: Century of growth in English education.* London: Longman, 1971.

DREW, C.J. Research on the psychological behavioral effects of the physical environment. *Review of Educational Research*, 1971, 41, 447-465.

ENGLEHARDT, D.F. *Aspects of spatial influence on science teaching methods.* Unpublished doctoral dissertation, Harvard University, 1968.

FINN, J.D. *A general model for multivariate analysis.* New York: Holt, Rinehart & Winston, 1974.

KANE, M.T. & BRENNAN, R.L. The generalizability of class means. *Review of Educational Research*, 1977, 47(1), 267-292.

KEEVES, J.P. *The home, the school, and educational achievement.* Unpublished Ph.D. thesis, Australian National University, Canberra, 1971.

NUNALLY, J.C. *Psychometric theory.* New York: McGraw Hill, 1967.

PROSHANSKY, H.M. Theoretical issues in environmental psychology. *School Review*, 1974, 82(4), 541-555.

PROSHANSKY, H.M., ITTLESON, W.H. & RIVLIN, L.G. *Environmental psychology: Man in his physical setting.* New York: Holt, Rinehart & Winston, 1970.

RYANS, D.G. *Characteristics of teaching.* Washington: American Council on Education, 1960.

SHAYCOFT, M.W. The statistical characteristics of school means. In J.G. Flanagan *et al.*, *Studies of the American high school.* Pennsylvania: University of Pittsburgh, 1962.

TRICKETT, E.M. & MOOS, R.H. Social environment of junior high and high school classrooms. In K. Marjoribanks (ed.), *Environments for learning.* London: National Foundation for Educational Research, 1974.

WEINSTIEN, C.S. Modifying student behaviour in an open classroom through changes in physical design. *American Educational Research Journal*, 1977, 14 (3), 249-262.

WILEY, D.E. Design and analysis of evaluation studies. In D.E. Wiley & M.C. Wittrock (eds), *The evaluation of instruction: Issues and problems.* New York: Holt, Rinehart & Winston, 1971.

Research in Science Education, 1978, 8, 110-118

THE ROLE OF LINKAGE MECHANISMS IN PROMOTING EDUCATIONAL CHANGE

John M. Owen

Introduction

In Australia, education has traditionally been administered by State Education Departments, and other agencies such as examination boards at the State or system level. Although there has been extensive devolution of authority for the conduct of many aspects of education to schools over the past decade, it is apparent that schools are still strongly affected by system level effects.

This paper is concerned with a comparison of the ways in which systems transferred information and gave support to a major curriculum innovation, the Australian Science Education Project (ASEP), the largest curriculum development project yet undertaken in this country. Because ASEP was a 'temporary system' (Miles, 1964), it had little opportunity to directly influence the way the materials it produced were used in schools (Owen, 1978). Thus agencies at the State level were primarily responsible for the introduction of the ideas and materials of ASEP to potential users.

Using a criterion set which measured aspects of the extent of penetration of ASEP materials into schools, this paper compares the strategies by which information about ASEP was transmitted by system level agencies to schools just prior to and during the first two years of the availability of the materials. A complete analysis of the roles of systems and the relationship between the administrative social and political factors which affected these roles is a necessary complement to the present paper, but time does not permit such an analysis here.

Theoretical considerations

There have been attempts to conceptualise the key mechanisms in the change process in social systems (Rogers & Shoemaker, 1971; Havelock, 1971; Schon, 1971). Research in educational diffusion and implementation have built on these foundations to identify the elements in curriculum change in educational systems (Owen, 1977) and the ways in which both dissemination (Rudduck & Kelly, 1976) and implementation (Fullen & Pomfret, 1977) have been enhanced.

These and other research findings were examined in order to identify the strategies by which knowledge of an innovation and its maintenance during the early stages of implementation were used by educational systems. The little research available suggested four major mechanisms were more important than others. These were:

(1) Participation in the formative stages of the program;
(2) Influence of formal teacher education programs;
(3) Written information, particularly in local publications; and
(4) Continuous "face to face" contact with linkage agents such as advisers.

It is important to note that these processes may serve to reinforce each other over the period in which the innovation was introduced, and that the presence of one in the absence of the others might limit the impact of the others. Nevertheless, in the first instance, it seemed appropriate to consider each of these mechanisms independently.

Data collection

In late 1975, a sample of approximately fifty schools in each state was drawn from the sampling frame of the Australian Council for Educational Research. Data which enabled estimation of the extent to which materials of the Australian Science Education Project had penetrated into Secondary Schools in each State was obtained from the coordinator of Science in these schools. In addition each teacher of Science provided information which included their knowledge of the characteristics of the ASEP materials, and their perceptions of importance various linkage mechanisms in transmitting about the ASEP development to teachers in schools. Overall a 90 per cent school response rate was achieved.

The information collected from these schools enabled conclusions to be made about the degree of acceptance of ASEP units in schools throughout the country (Owen, 1976). The analyses showed clearly that there were significant variation in the extent to which schools had obtained ASEP materials. There were also important variations between States. This suggested that there were factors present at the State level which influenced the diffusion and use of ASEP units. To investigate these effects further three States were chosen in which there were large differences in the average level of penetration of ASEP materials.

The States were:

(1) South Australia, where the impact of the ASEP materials was relatively high,
(2) Queensland, where the impact of the ASEP material was relatively low, and
(3) Victoria, where there was an average level of purchase between the two extremes.

For each of these States, a criterion set of measures were assembled which, collectively, described the penetration of ASEP materials into the schools, and these are presented in Table 1. The unit used to gauge the penetration of ASEP was the extent of purchase of module packs. A module pack consisted of 10 student textbooks. Two such packs would be adequate for a single class.

TABLE 1

Criterion set of measures of the impact of ASEP in three selected States

	Victoria (N=43)	*Queensland (N=48)*	*South Australia (N=42)*
Average school purchase[1] of module packs:			
to December 1974	12	3	18
January-October 1975	12	5	28
November 1975-February 1976	3	4	23
% schools with at least one module pack:			
up to December 1974	63	23	57
up to October 1975	81	40	88
up to February 1976	84	44	88
% schools in October 1975 with:			
0-19 module packs	22	19	10
20-80 module packs	54	19	64
>-80 module packs	5	2	14

1. Converted for school size. Values are per 500 Science students.

Victoria Most schools in this State purchased ASEP module packs and of these the majority first purchased in 1974, the year the ASEP materials became available. However, the majority of schools did not buy sufficient module packs to conduct a complete course using ASEP, as only five percent of schools bought 80 module packs or more by October 1975. From that date, interest in the materials waned, as only three percent of schools bought for the first time, and those schools which had already purchased greatly reduced their intake of ASEP units.

Queensland Committment to ASEP was generally low, both in the proportion of schools buying ASEP modules and in the extent of purchase. Average purchase of ASEP modules in the first two years of their availability remained at low levels, and was reduced in 1976. Only twenty-one percent of schools bought twenty modules or more.

South Australia Interest in ASEP built up over the first two years of its availability, and by October, 1975, nearly ninety percent of schools had purchased ASEP modules. Although no new schools bought ASEP in 1976, there were continued high purchases in that year. The generally high level of purchase was reflected in the factor that 78 percent of schools had bought twenty modules or more up to October 1975.

In summary, there were three different profiles. In Victoria, there was a relatively moderate purchase which fell away over time. In Queensland, there was a fairly constant low average purchase which varied little over time, and in South Australia, the reverse was true. These contrasting results formed the criterion set to examine the effects of system level strategies.

System level strategies

To gauge the effects of the system on the adoption and use of ASEP, information was collected from various sources. First, Heads of Science contributed data on a questionnaire concerning the effects of trials of the ASEP materials during the formation stages of the Project. Secondly, all Science teachers provided data regarding inservice education and contact with linkage agents. Thirdly, State level journals were scanned for articles relating to ASEP, to determine the number of such articles and their substance. Each of these strategies is now examined.

1. *Participation in the formative stages of the program*

In the ASEP development, the major form of teacher participation in the formative stages of the Project came about through 'national trials' of the materials, which were administered by each State Advisory Committee. There was also an opportunity for some schools to use the materials when they were recirculated after the formal trials concluded. Heads of Science in schools which had used ASEP were asked to comment on the influence of trials. Their responses are listed in Table 2.

TABLE 2

Effect of ASEP trials in schools on decision to adopt

Item: Indicate the relative importance of the following reason for ASEP materials being first purchased in the school.

'The Science department or those concerned with teaching Science were influenced by a favourable trial in the school'.

		Victoria (N=45)	*Queensland (N=46)*	*South Australia (N=43)*
Extreme	(5)	2	0	0
Major	(4)	11	7	10
Moderate	(3)	14	7	18
Slight	(2)	8	7	10
None	(1)	65	80	62
		100	100	100
	$\bar{x}$	1.77	1.43	1.76
	SD	1.17	1.18	1.07

Note: Figures in the main body of the table are percentages of all schools which used ASEP to October 1975.

The table shows that a large majority of schools were unaffected by trials of ASEP. A comparison of the results from the States shows that there was a greater impact of trials in Victoria than in the other two States. This was almost certainly due, not only to the effects of system level factors as defined, but also to direct contacts between the ASEP team and the schools doing 'local' rather than national trials.

An analysis of variance of this date showed significant differences between the States ($p < 0.01$). The results of a multirange test, attributed to Scheffe (1959) showed significant differences between the effects of trials in South Australia and Victoria on one hand, and Queensland on the other.

2. *Formal teacher education programs*

Teachers were asked to indicate the number of occasions in which they had attended inservice conferences on ASEP during the period 1974-1975, the years in which responsibility for the implementation of ASEP materials became that of agents based at the State level. The results are indicated in Table 3.

It is clear that the vast majority of teachers of Science never attended an inservice course related to ASEP during this period. There was a similar pattern in Queensland and South Australia, where about 85% of teachers never attended inservice courses. In Victoria, one fifth of all teachers reported that they had attended a one-day conference. A series of conferences were organised by the ASEP State Advisory Committee in Victoria in conjunction with the Science Teachers Association of Victoria in this period and it is obvious that these were reasonably well attended by Victorian teachers.

An analysis of variance carried out on these data showed that there were no significant differences between the States, and it can be inferred that attendance at inservice activities could not explain the differences in the impact of ASEP across the States.

TABLE 3

Total number of days spent at an ASEP inservice course, 1974–1975
(Teacher responses)

		Victoria (N=234)	*Queensland (N=329)*	*South Australia (N=288)*
Four or more	(5)	2	4	4
Three	(4)	2	2	2
Two	(3)	7	4	3
One	(2)	20	6	6
Zero	(1)	68	84	85
		100	100	100
	$\bar{x}$	1.47	1.36	1.34
	SD	0.89	0.93	0.96

Note: Figures in the main body of the table are percentages.

Another possible linkage was through preservice teacher education programs.[1] It was possible that the preservice experiences of beginning teachers could have influenced the decisions by schools to use ASEP materials. To examine the extent to which the ideas and materials of ASEP development were emphasised in preservice teacher-education programs, information was collected from teachers in the sample who had recently begun working in schools. These are recorded in Table 4.

An analysis of variance of the results in Table 4 showed that significant differences occurred between the States ($p < 0.01$). The Scheffe Test showed that two homogeneous subsets were present, with Queensland and Victoria in one group and South Australia in the other.

This was an interesting result because the rank order of the press for ASEP from College and University education faculties between States was the reverse of the rank order of the extent of impact in terms of sales of materials and number of schools affected. In Queensland there was a strong emphasis on the ASEP development in these institutions, whilst in South Australia, almost one-half of all beginning Science teachers in the period 1973–1975 entered schools without being aware of the range of curriculum materials available for use in Science classes. This suggested that up to 1975 at least, one or more of the major teaching training institutions in South Australia did not regard the study of the rationale underlying curriculum development in Science and the characteristics of the curriculum materials available as an essential element in the preparation of prospective Science teachers.

3. *Written information* (Journals)

There is some evidence, from studies in other countries that journals, particularly those of local teachers' associations were considered to be a major means by which teachers obtained information about a curriculum project (Yegge, 1971).

---

[1] Rosier (1973) has shown that Science teachers in Australia are, on average much younger than their counterparts in other countries. It is possible, at least in some Australian States that a teacher with three years teaching experience would be a Head of Science in a school.

TABLE 4

Preservice education emphasis on ASEP

Item: To what extent did your University or College classes in your preparation to be a Science teacher refer to ASEP materials?

		Victoria (N=89)	*Queensland (N=129)*	*South Australia (N=128)*
A range of curriculum materials discussed, but ASEP given more prominence	(4)	19	33	17
A range of curriculum materials discussed, but ASEP not given special prominence	(3)	48	36	30
A range of curriculum materials discussed, but ASEP was not among them	(2)	13	15	9
There was no specific discussion of curriculum materials for science in my teacher training	(1)	19	16	43
		100	100	100
	$\bar{x}$	2.65	2.86	2.19
	SD	1.00	1.05	1.22
% teachers responding (beginning in 1973, 1974 or 1975)		38	39	44

Note: Figures in main body of the table are percentages.

In Australia, the geographical location of schools even within States of smaller size suggests that teachers' journals should be one of the most important vehicles for communication among teachers. State Science teachers' journals in Australia are produced by State Science Teachers Associations and tend to be important to teachers because of the number of articles devoted to local rather than national aspects of Science teaching.

There are no current figures available on the proportion of all teachers of Science who are members of the Science Teachers' Associations; the only reliable figures published were those by Rosier (1973) who reported that the percentage of all teachers of the Sciences who were members of a professional subject-matter teachers' association in 1970 were as follows: 37 percent in Victoria, 41 percent in South Australia and 48 percent in Queensland. Thus, if these values held during the next four years, between four and five teachers of Science in every ten received journals produced by the State Associations.

However, it is important to note that during the period 1971 to 1975 there were differences in the number of issues of the journals between States. While *Lab Talk* (Victoria) and *SASTA Journal* (South Australia) were published regularly during this period, *The Queensland Science Teacher* (Queensland) was not published, due to financial difficulties from 1973 to 1975. Consequently, there was less chance for activities relating to ASEP to be published in *The Queensland Science Teacher*. To obtain an indication of the potential of these three journals to transmit information about ASEP to practicing teachers, a count was made of articles devoted to ASEP between 1971 and 1975. There were 24 articles in *Lab Talk*, nine in *SASTA Journal* and three in *The Queensland Science Teacher*. Clearly, teachers in

Victoria were in a position to follow the development of ASEP more closely than teachers in the other two States. In Queensland, it was apparent that information about ASEP was not readily obtainable from the local journal. Finally, it should be noted that all members of the State Associations receive *The Australian Science Teachers' Journal* (ASTJ). Articles about ASEP were included in ASTJ at regular intervals during this period. However, because it is a national level journal, ASTJ often tends to focus on broad issues with a bias toward educational research, rather than on issues related to the solution of day-to-day problems of adopting and using a new curriculum.

While an analysis of this nature needs to be considered with some care, it appears that there were probably differences in the influence of the three journals on teachers of Science in the three States.

4. Advisory personnel

In addition to inspectors of schools, advisers or consultants have been used by Education Departments in recent years to assist teachers in the conduct of their courses. Table 5 records the opinions of teachers about the extent of support given by these officers to the ASEP development.

TABLE 5

Teachers perceptions of science inspectors/consultants attitudes to ASEP

		Victoria (N=288)	*Queensland (N=319)*	*South Australia (N=285)*
More emphasis on ASEP than on other Science materials	(4)	11	2	17
No marked emphasis on ASEP	(3)	13	25	28
Attitude of inspectors/ consultants not known	(2)	77	67	55
Other materials rather than ASEP promoted	(1)	0	6	1
		100	100	100
	x	2.37	2.23	2.63
	SD	0.67	0.65	0.77

Note; Figures in body of table are percentages.

The most striking result in Table 5 is the high proportion of Science teachers who had no knowledge of the attitudes of these advisers to the ASEP development materials. A one-way analysis of variance applied to the figures showed that the differences between States were significant ($p < 0.01$). Teachers in South Australia reported significantly greater support for ASEP from advisers than did teachers in the two other States.

Table 5 also showed that, of the 33 percent of responses which implied a knowledge of the attitudes of curriculum materials in Queensland, about one fifth reported that there was promotion by advisers of curriculum materials rather than ASEP. It would appear that there was an underlying objection to the widespread use of ASEP by at least some of the advisers in this State. This was an intriguing result in view of the fact that the outline syllabus in

Queensland has a set of aims which are strikingly similar to those underlying the ASEP development.

TABLE 6

Linkage mechanisms – summary

Criterion	*Ranking*
1. participation in trials	$\boxed{\text{QLD}}$ $\boxed{\text{SA} \quad \text{VIC}}$ →
2. (a) inservice education	$\boxed{\text{SA} \quad \text{QLD} \quad \text{VIC}}$ →
(b) preservice education	$\boxed{\text{SA}}$ $\boxed{\text{VIC} \quad \text{QLD}}$ →
3. local journal articles	$\boxed{\text{QLD}}$ $\boxed{\text{SA}}$ $\boxed{\text{VIC}}$ →
4. advisory personnel	$\boxed{\text{QLD} \quad \text{VIC}}$ $\boxed{\text{SA}}$ →

Note: Boxes show homogeneous sets of subgroups according to the Scheffe test. The exception is the criterion 'Local Journal Articles'. In this case, the States are ranked in order of the number of articles found in the journals.

Conclusion

In general, there were fewer linkages regarding ASEP to teachers of Science in Queensland than to their counterparts in the other two States. The exception to this generalisation was in preservice education. The negative correlation between extent of purchase of ASEP and the preservice press for the materials suggested that beginning teachers, those in their first, second and third year of teaching had little influence on the planning of Science courses in the schools in which they first taught.

In South Australia, high and continuing purchases of ASEP materials occurred in the period 1974–1976. From the profile of possible linkage mechanisms, contacts of teachers with consultants was notable. This suggested that there was a connection between the activities of these consultants and the continuing demand for ASEP in that State. The maintenance of the innovation during this period appeared to be associated with a resource system based on personal interaction between expert consultants and users. The availability of a source interaction network with a 'permanent system' the Education Department of South Australia was obviously of some importance in meeting the need for assisting in curriculum changes initiated by school Science departments.

Despite warnings from those in the ASEP development team that support for teachers was essential during the implementation stage (Ramsey, 1970) many teachers were not directly affected by linkage mechanisms designed to introduce the ideas and materials of the ASEP development. Observers such as Dynan (1978) have noted that lack of implementation support for other large scale projects have sounded their death knell. Will the sound of the future for the ideas and materials of ASEP be only a muffled drum?

References

DYNAN, M.B. Aspects of curriculum innovation in science education. *Unicorn*, 1978, 4 (1), 23-32.

FULLEN, M. & POMFRET, A. Research on curriculum instruction and instruction implementation. *Review of Educational Research*, 1977, 47 (2), 335-397.

HAVELOCK, R.G. The utilization of educational research and development. *British Journal of Educational Technology*, 1971, 2, 84-98.

MILES, M.B. On temporary systems. In M.B. Miles (ed.), *Innovation in education*. New York: Teachers College, Columbia University, 1964, 437-492.

OWEN, J.M. Three years on: The impact of ASEP materials in Australian schools. *Australian Science Teachers Journal*, 1975, 21 (3), 15-24.

OWEN, J.M. School factors affecting curriculum change in science. *Research in Science Education*, 1977, 7, 61-70.

OWEN, J.M. *The impact of the Australian science education project on schools.* Canberra: Curriculum Development Centre, 1978 (in press).

RAMSEY, G.A. Paper given at the Guidelines Conference of the Australian Science Education Project. Monash University, January, 1970.

ROGERS, E. & SHOEMAKER, F.F. *Communication of innovations* (2nd ed.). New York and London: Free Press and Collier, MacMillan Ltd., 1971.

ROSIER, M.J. A statistical picture of Australian science teachers. *Australian Science Teachers Journal*, 1973, 19 (4), 27-39.

RUDDUCK, J. & KELLY, P. *The dissemination of curriculum development.* London: National Foundation for Educational Research, 1976.

SCHON, D.A. *Beyond the stable state. Public and private learning in a changing society.* Harmondsworth: Penguin, 1971.

YEGGE, J.F. The adoption of an innovation in physics teaching. Unpublished Ph.D. thesis. ERIC Microfiche ED52 961, 1971.

Research in Science Education, 1978, 8, 119-128

SOME OUTCOMES OF AN ENVIRONMENTAL KNOWLEDGE AND ATTITUDES SURVEY IN ENGLAND

James M. Richmond

Introduction

In recent years there has been an increasing demand in England for school courses and examination syllabuses in environmental education. This has resulted in extensive efforts in curriculum development at both the local and national levels, including the Schools Council projects "Environmental Studies" and "Project Environment" (Schools Council, 1976).

Unfortunately the educators involved in these curriculum developments have had little objective data at hand regarding the current environmental knowledge and attitudes of their student populations. They have therefore been forced to work by trial and error and have tended to be subjective in deciding the content and methods most suitable for their programmes.

With this in mind, a national survey was conducted in England in 1976 to establish baseline data relating to the environmental knowledge and attitudes of 5th year secondary students. The pattern of student responses and an examination of the relationships between environmental knowledge and attitudes and the variables of sex, school type, sex composition of the school, school size and region have been presented in a report elsewhere (Richmond & Morgan, 1977) and will not be discussed in any detail here. This paper will focus on several outcomes that emerged from the data. These outcomes relate to:

(1) the relationships between factual knowledge, conceptual knowledge and attitudes;
(2) student perceptions of the most serious environmental problems both locally and nationally; and
(3) student perceptions of the major source of their environmental knowledge.

Design of the study

The Instrument

The instrument developed for the survey consisted of three questionnaires, Forms A, B and C. Part 1 of each form contained factual knowledge and perceptual questions, Part 2 dealt with conceptual knowledge, and Part 3 presented 15 statements of belief for pupil reaction. There were a total of 45 items on each form, with 14 common items providing the means for comparing response patterns on the three questionnaires. Figure 1 shows the model used in developing the instrument. Knowledge and belief questions relating to each of the identified "environmental concern" categories were selected from a pool of almost 400 items to fill every cell in the matrix.

The pilot instrument was critically analysed by 18 educators and tested in a pilot study in England involving 386 students in nine schools. Based upon these analyses, appropriate changes were made in producing the final instrument. Test/retest procedures involving 164 students were later used to determine the reliability of the instrument. The reliability coefficients were 0.84, 0.83 and 0.89 for Forms A, B and C respectively.

	COGNITIVE *Measure of Environmental Knowledge*		*AFFECTIVE* *Measure of Attitude toward the Environment*
Environmental Concerns	*Factual Questions*	*Conceptual Questions*	*Belief Questions*
Pollution			
Population			
Natural Resources			
Land Use			
Energy			
Env. Health/Safety			
Ecological Relationships			
Soc./Pol./Ec. Influences			
	17 Questions per Form Answer Format: Multiple Choice	10 Questions per Form Answer Format: True/False/Don't Know	15 Questions per Form Answer Format: Agree/Disagree/No Opinion

FIGURE 1

Matrix defining environmental concerns and type of questions included in the instrument

The sample

The target population consisted of all 5th year students enrolled in the secondary schools of England. The 5th year was chosen since it represents the last year of formal schooling for a large proportion of secondary students.

A sampling procedure was used that would ensure proportional representation of the major types of school (viz. comprehensive, secondary modern, grammar, direct grant, independent, and "other secondary") in every region of the country. A total of 500 secondary schools was selected in the sample, and within each participating school the instrument was administered to a sub-sample of about 30 students in the 5th year.

Data collecting procedures

It was decided that the most effective method for collecting data would be to mail the testing materials directly to the 500 schools selected in the sample, with a carefully worded letter of explanation to the head teacher. A total of 383 schools (or 76.6% of the sample) returned completed answer sheets, providing information from 11 009 students. This exceptionally high response rate for a national survey probably resulted from employing a number of procedures and techniques designed to encourage co-operation. Many of these procedures were based upon Eyer's study (Eyers, 1977) and further details may be found in the publication by Richmond and Morgan (1977, p. 32-33).

Results

Measures of environmental knowledge and attitudes

In general, students responded poorly to the factual knowledge questions. The overall correct response rate was approximately 46%, with only 14 of the 43 factual knowledge items being correctly answered by more than 50% of the respondents.

Students demonstrated a greater understanding of environmental concepts, with an overall correct response rate of a little over 60%. Seventeen of the 24 conceptual knowledge items were correctly answered by more than 50% of the students.

A moderately positive attitude toward the environment was indicated by the response pattern on belief items. About 60% of all belief responses were environmentally positive (in agreement with the panel), and on 27 of the 37 items more than 50% of the students selected the environmentally positive alternative. However it should be noted that the students' expressed environmental attitudes tend to be strongly positive when the object of concern does not impinge directly on their lives, but appear to be relatively weak when some personal sacrifice is required. For example, over 84% of the respondents agreed that "man has a moral responsibility to protect the natural environment" while fewer than 45% agreed that community standards for pollution are more important than industrial growth and development, or that most couples should not produce more than two children.

These results are remarkably similar to the response patterns observed in similar studies conducted by Perkes (1973) and Bohl (1976) in the United States and by Eyers (1975) in Australia. Some areas of concern do emerge from the common findings. Firstly, these studies have revealed a number of misconceptions about the environment and an overall poor grasp of factual information. Without a firm foundation of knowledge one must question the ability of the group as a whole to make responsible decisions regarding their environment when they assume voting rights and positions of authority in the near future.

Secondly, the apparently positive attitudes measured by the instrument should give little cause for complacency. As noted above, positive responses were elicited by statements that did not involve personal commitment while more negative attitudes were revealed when personal sacrifice is required. Thus one must question whether the expressed attitudes are deeply rooted and based upon knowledge, experience and conviction, or whether they are superficially "learned" or instilled by indoctrination and will not translate into responsible behaviour.

Relationships between factual knowledge, conceptual knowledge and attitudes

Since one of the frequently stated aims of environmental education programmes is to develop positive attitudes toward the environment, a fruitful area for the environmental educator to explore may be in the contentious field of attitude development. After an extensive search of the literature, Bohl (1976) reported that there was consistent support of a direct relationship between knowledge and attitude. In his own study he found that "Low correlations were characteristic between factual type cognitive variables and belief type attitude variables" while "Significant correlations were noted between conceptual type cognitive variables and belief type attitude variables" (Bohl, 1976, p. 162).

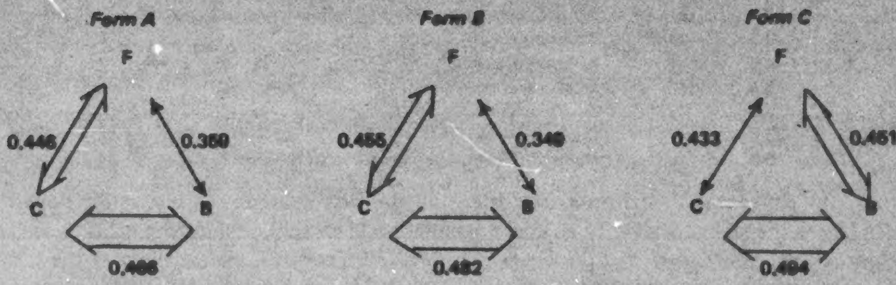
Rhine claimed that "there is already an established theoretical link between concepts and attitudes" and defined an attitude as "a concept with an evaluative dimension" (Rhine, 1967). However no research data was evident to the author to indicate the strength of this theoretical link between concepts and attitudes. With this in mind, the survey instrument was designed so that relationships between factual knowledge, conceptual knowledge and attitudes might easily be measured (see Fig. 1). Correlation coefficients were computed between the total scores on the factual, conceptual and belief sections of each form using SPSS sub-program SCATTERGRAM, and SPSS sub-program PEARSON CORR was used to determine the Pearson product-moment correlation coefficients between all items on Forms A, B and C.

The strength of the relationships between the total scores on the factual, conceptual and belief sections of each form is represented diagramatically in Figure 2. The correlation coefficients (all significant at less than the 0.001 level) are shown adjacent to each arrow.

The Hotelling t-test (for comparison of two correlation coefficients) shows that in each case the correlation between conceptual knowledge and beliefs is greater than between factual knowledge and beliefs at the 0.01 level of significance.

That the strongest relationship exists between conceptual knowledge and attitude, while the weakest relationship is between factual knowledge and attitude, is supported by examining the correlations between individual items. Significant correlations ($r > 0.10$) were found to exist between 36.9% of the conceptual and belief items, 23.3% of the factual and conceptual items, and 15.0% of the factual and belief items.

These results provide a more precise measure of the strengths of these relationships than any of the previous studies concerning environmental knowledge and attitudes. Without diminishing the value of factual environmental knowledge (which was mentioned earlier as a prerequisite for responsible decision-making), these results appear to underline the importance of conceptual knowledge in the development of positive environmental attitudes. Although



F = Factual knowledge scores; C = Conceptual knowledge scores; B = Belief scores (attitude)

FIGURE 2

Correlations between factual, conceptual and belief scores

no causative relationship has been demonstrated, the relatively strong correlation between the conceptual and belief sections suggests that the development of sound concepts might be a productive means of leading to the establishment of positive attitudes. The importance of conceptual development has been stressed by many educators, and these findings not only support their position but are a reminder that conceptual understanding should be a prime objective of environmental education programmes.

Student perceptions of environmental problems

In an attempt to determine which environmental problems were considered by secondary students to be most serious in both their home community and in the nation as a whole, two "perceptual" questions were included in the inventory. These questions are stated below with the percent response indicated beside each alternative.

Which *one* of the following problems do you think is *the most serious* in the community where you live?

14.4	a)	Land use
12.2	b)	Traffic accidents
10.4	c)	Air pollution
3.5	d)	Water pollution
11.3	e)	Rubbish disposal
8.2	f)	Over-crowding
4.1	g)	Public health
14.5	h)	Crime
16.3	i)	None of the above are problems in our community.

An unexpected outcome on this question was the fact that the largest group of respondents did not perceive any of the listed problems to be applicable to their immediate surroundings. The societal problem of "crime" was selected more frequently than any of the physical environmental concerns as being most serious in the local communities.

Which *one* of the following problems do you think is *the most serious* in Britain?

9.1	a)	Land use
9.4	b)	Traffic accidents

12.2	c)	Air pollution
8.3	d)	Water pollution
6.6	e)	Rubbish disposal
26.4	f)	Over-crowding
5.2	g)	Public health
22.0	h)	Crime
0.9	i)	None of the above are problems in Britain.

In contrast with the response pattern on the first question, all but a few students (0.9%) were prepared to identify an environmental problem applicable to the national scene. Over-crowding, which was of little concern in local communities, was selected as the most serious environmental problem for Britain as a whole. This was closely followed by "crime" indicating that these societal problems loom larger in the minds of young people than problems relating to the physical environment such as air and water pollution.

The fact that an appreciable number of students recognize that their country is afflicted with environmental problems, but do not recognize that these problems occur in their home communities, points to a serious lack of awareness of their immediate surroundings and may indicate the need for an increased emphasis on local studies.

Student perceptions of source of environmental knowledge

For his survey conducted in Australia in 1974, Eyers devised a question "to provide information about the sources of environmental knowledge, as perceived by the students themselves" (Eyers, 1975, p. 72). The same item was included on all forms used in the survey of English students. The response rates are listed below beside each alternative on the question.

Which *one* of the following best describes the way in which you have gained most of your knowledge about the environment?

31.5	a)	general education at school
6.9	b)	special environmental courses at school
48.1	c)	private reading, the radio, and TV
13.5	d)	talking with parents, friends and other people.

It is interesting to note that fewer than 40% of the students believed that they had gained most of their environmental knowledge from learning at school, while over 60% indicated that this knowledge had been gained from activities that might be described as "self-education". The media appears to have played a major role in providing environmental knowledge while special education courses have made a relatively small impact.

An additional interesting observation is that those students who identified the media as the major source of their knowledge scored significantly higher than the other groups on factual knowledge, conceptual knowledge and beliefs; while those who selected "special environmental courses at school" scored significantly lower than the other groups on all measures.

Since this item was designed to elicit the students' perceptions of where they have gained most of their environmental knowledge, and does not necessarily indicate the true source, some caution should be observed in interpreting the response pattern. For example, the fact that fewer than one-half of the students believe that they have gained most of their knowledge in the classroom does not necessarily imply that schools are not doing an adequate job in environmental education; however it does tend to raise that suspicion. Perhaps the most interesting outcome from this question is the importance attributed by students to the media as a source of environmental knowledge. While attitudes developed through the passive reception of information may not have the same influence on behaviour as attitudes arising from personal investigation and involvement, it nevertheless seems wise to look afresh at the potential of the media in this area. In addition to improving the quality and quantity of environmental education in the school curriculum, it would appear to be a fruitful strategy to intensify the coverage of environmental matters in newspapers and on the radio and television.

References

BOHL, W.B. A survey of cognitive and affective components of selected environmentally related attitudes of tenth and twelfth grade students in six midwestern, four southwestern and twelve plains and mountain states. Unpublished Ph.D. dissertation, The Ohio State University, 1976.

EYERS, V.G. Environmental knowledge and beliefs among grade 10 students in Australia. Unpublished Ph.D. dissertation, Oregon State University, 1975.

EYERS, V.G. They don't know much but their hearts are in the right place (so they say). Paper delivered at the ASERA Conference, Wagga Wagga, N.S.W., 1977.

PERKES, A.C. A survey of environmental knowledge and attitudes of tenth and twelfth grade students from five great lakes and six far western states. Unpublished Ph.D. dissertation, The Ohio State University, 1973.

RHINE, R.J. A concept-formation approach to attitude acquisition. In M. Fishbein (ed.), *Readings in attitude theory and measurement.* New York: John Wiley & Sons, 1967, 382-383.

RICHMOND, J.M. & MORGAN, R.F. *A national survey of environmental knowledge and attitudes of fifth year pupils in England.* Columbus, Ohio: Educational Resources Information Center, 1977.

SCHOOLS COUNCIL. Project profiles (6th ed.). London: Schools Council, 1976.

Research in Science Education, 1978, 8, 127-134

AN EXPLORATORY STUDY IN SCHOOL ASSESSMENT: THE RELATIONSHIP BETWEEN ACHIEVEMENT AND ATTITUDE SCALES

M. N. Maddock

The problem of school based assessment and the affective domain

Over the years there has been a greatly increased emphasis on school-based curriculum development, and with the decline in the emphasis on external examinations, there has also been increasing responsibility placed on the classroom teacher to develop and apply evaluation and assessment procedures which provide a level of comparability of results across schools.

The demise of the external school certificate examination at fourth year level in N.S.W. is a relatively recent phenomenon, compared with elsewhere in Australia, and it has brought with it considerable difficulties for the classroom teacher. The fading out of the external examination system has coincided with the issuing of a science syllabus which consists of a set of aims and objectives, and a vague statement of content, which is merely a short list of general headings (Secondary Schools Board, 1973a).

Today's classroom teacher is thus faced with considerable responsibility in selecting and organising content and also in carrying out the student assessment programme. Testing has always been part of a teacher's repertoire, but never has such a high degree of responsibility rested on his shoulders to ensure high levels of reliability, validity and also comparability with testing in other schools.

The issue has been complicated in N.S.W. because of the insistence by the Schools Board that affective domain objectives be evaluated and the results included in the assessment programme. In the statement of the Syllabus in Science Years 7-10 (Secondary Schools Board, 1973a), the document clearly spells out objectives in the area of interests and attitudes.

The Assessment Guide, issued by the Secondary Schools Board (1973b, p. 2) states:-

> The Junior Science Course is considered to be concerned with three main areas of pupil development (see *Syllabus Aims*) – INTERESTS and ATTITUDES, KNOWLEDGE and UNDERSTANDING and SKILLS. When accumulating details of progressive achievement in order to produce the final assessment, teachers are advised to spread the available range of marks according to:
>
> | | |
> |---|---|
> | Interests and Attitudes | 0–30 |
> | Knowledge and Understanding | 0–50 |
> | Skills | 0–20 |

This means that teachers have to devise techniques for evaluating interests and attitudes, and then convert the results to a component of an overall assessment, a task of considerable difficulty considering the relatively primitive state of the technology of instrumentation in the affective domain, and the limited knowledge of test theory and techniques held by many teachers.

The freedom of choice open to teachers, coupled with the historical social demands for uniformity of testing standards across the state of N.S.W., has created something of a dilemma for the teacher. Preliminary research by Telfer (1977) at the University of Newcastle indicates that beginning teachers place assessment as one of their highest ranked problem areas. In a sample of 106 teachers in their first year during 1977, pupil assessment was ranked as the foremost professional difficulty. It was ranked number one problem by 32% of the sample, ahead of programming (27%), curriculum and course construction (20%) and teaching method and style (15%).

McBurney (1978), in a survey of the effectiveness of implementation of the new syllabus, reports strong polarisation concerning the place of attitude assessment in a sample of over 300 science teachers. Although 89% agreed that some statement of attitudes was required, and 87% agreed that the present aims represented a satisfactory or better statement, 60% wanted some form of modification. Forty three percent were in favour of assessing attitudes for certification, but 44% were against it. Seventy three percent favoured assessment for feedback purposes, 63% for diagnostic purposes, but only 17% stated that attitudes should not be assessed in any form at all.

This difficulty has been recognized by the Diploma in Education course at the University of Newcastle, which has included an elective in testing and evaluation in its Group 1 offerings of foundational units, and where a substantial section has also been included in the Science Curriculum and Method unit, for trainee science teachers.

Developing and attitude scale

During 1977, the students enrolled in the Science Curriculum and Method Unit were given the task of undertaking a cooperative, class project to develop a scale aimed at evaluating one of the affective domain objectives stated in the science syllabus.

The class chose to attempt to develop an attitude instrument aimed at the objective

to develop a belief in the value of the conservation of our natural environment

which is listed in the Statement of Aims and Content (Secondary Schools Board, 1973a, p. 11). They decided to use the item-building technique in which a concrete statement of a situation is stated and three alternative modes of thought or action, representing the low, middle and high end of a continuum, are posed, similar to that developed by Maddock (1974).

The responses were weighted on a 3 point scale, with the maximum weight being allocated to a behaviour description defined as representative of the upper end of the attitude continuum, the minimum to one defined as representing the lower end, and the middle value to one representing some intermediate position.

By working in groups, the students produced a pool of items which were then handed over to an editing committee. This committee selected 22 of the items, which were considered the best ones, and further refined these to produce the final scale. The items covered a wide range of environmental conservation subjects such as the use of pressure-pack sprays, packaging, road building and nature reserves, recycling of resources, wild life, and the exploration and recreational use of bushland.

The following example is representative of the types of items included in the scale:

It has been found that certain brands of spray cans contain chemicals which break down ozone. Ozone in the upper atmosphere prevents harmful radiation reaching earth. If you were buying fly spray for your mother, would you

1. Conscientiously select a brand that doesn't contain these chemicals?
2. Buy the brand mum uses, regardless of its content?
3. Buy the brand you have found works best?

Relationship between attitude and cognitive scales

Apart from the difficulties involved in developing scales which attempt to make measurements in the affective domain, there is the problem of what relationship exists between scores obtained on scales of this type and scores obtained on normal cognitive tests. The validity of adding scores obtained in attitude and interest measures to those obtained in assessment of cognitive objectives and using the result in any ranking of students exercises depends on the sum being a representative measure of some uniform attribute.

To obtain any meaningful value, the scales used must be closely correlated, otherwise they would be best reported as separate measures. In addition, the scales used must have high levels of reliability, to minimize measurement error.

In this project, apart from attempting to develop an attitude scale, the purpose was to also produce cognitive, achievement tests, to analyse the results of these tests to assess the reliability of the scales produced and to investigate the relationship between the scores obtained on the cognitive scales and those achieved on the attitude scales, as an indicator of the validity of obtaining any additive measure.

Investigating relationships

Depending on the classes taught during the six week practice teaching in term 2, the students worked, either individually or in small groups, to produce 13 separate cognitive tests, which were administered to 20 classes at 6 Newcastle area high schools (464 pupils in all), 5 classes at grade 11 level, 6 at grade 10, 4 at grade 8 and 5 at grade 7. In some cases, where a common test was administered across all classes in a single grade in a school, analyses were carried out on individual classes as well as on the combined results. Where possible, the results of the cognitive tests were correlated with the results of the attitude test for each set of results obtained, 15 correlations being obtained in all.

The tests were administered during third term to the classes they had taught for block practice teaching in second term. Each student prepared his or her own set of results for computer processing. For each set of results, printouts were obtained with listings of descriptive statistics, item analyses data, and product moment correlations between tests.

Investigating grade level effects on the attitude scale scores

When a specific objective is explicitly stated in a syllabus document, with the intention that teaching be directed towards its achievement, it would be expected to find progressive rise across grades in the level of scores on any measure aimed at evaluating the objective, if the teaching programme is in any way successful.

In the case of an objective such as to develop a belief in the value of conservation, of course the school is not the only influence at work, and this study made no attempt to control for other variables. However, the expectation of a rise in scores across grades is still a reasonable one.

To investigate if any grade level effect was present and also to investigate whether any difference existed between boys and girls, a 4 x 2 analysis of variance was carried out, using the boy-girl dimension of sex and grades 7, 8, 10 and 11 over the length of high schooling dimension.

Performance of the attitude scale

For a first trial in an instrument development project, the attitude scale performed at a very satisfactory level. The coefficient alpha reliability for the total sample of 464 subjects was 0.56, of an order which Nunnally (1967) regards as suitable for early stages of research. Seventeen of the items produced correlations with the total score of +0.25 or better, and only two items were under +0.20. Choices were distributed well across all three alternatives.

The cognitive scales

Twenty three sets of results were obtained from the twelve cognitive tests set, with a very wide range of coefficient alpha reliabilities being obtained. No test produced a coefficient alpha of the order of 0.85 or above, as suggested by Nunnally (1967) as necessary for discriminating between individuals, two of the sets of results provided alphas of 0.75 or above (0.75 and 0.80 respectively) and 10 had alphas below the 0.60 level indicated by Nunnally (1967) as a suitable one for preliminary research. The obtained alphas are summarized in Table 1.

Correlation between cognitive scales and the attitude scale

All the fifteen correlations obtained between the scores on cognitive scales and the scores on the attitude scale were low, ranging from −0.34 in one small grade 11 class to +0.37 on another small grade 11 class. However the small class size and the low alpha reliabilities obtained on some of these, render interpretation very tenuous. However, seven sets of results were obtained where class size and the level of coefficient alpha were considered sufficient for reasonable interpretation to be made.

The correlations obtained in these seven cases were as follows:

+0.07, +0.90, +0.10 (2), +0.11, +0.21* and +0.32*

The two correlations marked with an asterisk were statistically significantly different from zero at the 0.05 level. The evidence therefore supports the conclusion that there is little relationship between the attitude, as measured by the scale developed, and cognitive achievement in science at high school, as measured by this set of tests. This of course, seriously raises a challenge to the validity of using the scores on this scale with scores on cognitive scales, to produce an additive results. The results obtained are summarized in Table 1.

Grade level and sex effect

The analysis of variance model used was as follows, where S represents sex and E represents grade level:

$$Y = M + S + E + SE + ERROR$$

Grade levels investigated were 7, 8, 10 and 11. The design, means and standard deviations obtained are summarized in Table 2, and the results of the analysis of variance are summarized in Table 3.

There was no significant difference between the mean for boys (49.9) and the mean for girls (51.1), the F value being 2.72, nor was there any sex-grade level interaction effect (F=0.45).

TABLE 4

Reliabilities and Intercorrelations
Cognitive Tests and Attitude Test

School	*Grade*	α *Cognitive*	α *Attitude*	*Cognitive-Attitude*	*School*	*Grade*	α *Cognitive*	α *Attitude*	*Cognitive-Attitude*
1	10	0.51	0.67	0.11	4	8 (1)	0.76	0.41	
	11 (1)	0.26	0.07			8 (2)	0.71	0.58	
	11 (2)	0.60	0.16	0.28**		combined	0.73	0.47	0.32*
	11 (3)	0.58	0.29	-0.34**					
	combined		0.55						
2	8 (1)	-0.13	0.43	0.21**	5	7 (1)	0.58	0.22	-0.07**
	8 (2)	0.29	0.43	0.34**		7 (2)	0.40		
	9 (3)	0.37	0.39			7 (3)	0.49	0.60	0.09
	11	0.33	-0.04	0.37**		7 (4)	0.48	0.57	0.10
	combined		0.55			7 (5)	0.52	0.34	0.10
						combined		0.49	0.07
3	10	0.36	0.33	-0.16**	6	10 (1)	0.81	0.47	
	11	0.26	0.72			10 (2)	0.54	0.58	
	combined		0.30			10 (3)	0.47	0.82	
						10 (4)	0.66	0.75	
						combined	0.71	0.65	0.21*
2-3		0.63		+0.24	combined schools			0.56	

**Class sizes too small or coefficient alpha reliabilities too low to be safely interpreted.

*$p < 0.05$

TABLE 2

Design
Analysis of variance
Sex by grade level

	Boys (S1)			*Girls (S2)*			*Total*
	N	*Mean*	*SD*	*N*	*Mean*	*SD*	*Mean*
Grade 11 (E1)	16	5.24	3.5	51	52.8	4.5	52.8
Grade 10 (E2)	98	50.6	5.6	88	51.9	4.8	51.2
Grade 8 (E3)	55	48.3	4.8	59	48.6	4.7	48.5
Grade 7 (E4)	43	49.5	5.4	54	51.0	3.7	50.3
Total (Mean)		49.9			51.0		

TABLE 3

Summary of analysis of variance

Effect	*S.S.*	*F*	*df F*	*P<*	*E*	*E^2*
Sex	63.99	2.72	1/456			
Education	771.83	10.94	3/456	0.001	0.247	0.061
Education-Sex interaction	31.55	0.45	3/456			
Error	10728.21					

However the grade-level effect was significant at the 0.001 level, with an obtained F of 5.42. The epsilon calculated from this F value (Cohen, 1965) of 0.25 indicates that about 6 percent of the variance is associated with grade level membership.

Scheffé tests, used to determine the significance of the difference between specific means (Kerlinger & Pedhazur, 1973), showed that the mean for grades 10 and 11 combined is significantly higher than the mean for grades 7 and 8 combined, and that the mean for grade 10 is significantly greater than that for grade 8. However, there was no significant difference between the grades 7 and 8 means, or the grades 10 and 11 means. The results of the Scheffé test are summarized in Table 4.

In the sample investigated, there was a grade level effect in the direction that could be predicted (except that the lowest mean was grade 8), assuming that science teaching in schools is in any way geared to achieve the stated objective of improving attitudes towards conservation, however as indicated earlier, caution must be experienced in interpreting this result, because no attempt was made to control for other variables. However, the magnitude of the effect is only a moderate one, the two significant differences being 0.4 σ between junior and senior grades and 0.5 σ between grade 7 and grade 8. It could be tentatively concluded that the effect of five years of high school science teaching is not really a major one in developing the belief in the value of conservation of the environment.

Conclusion

Increased emphasis on school based curriculum development and assessment, with

TABLE 4

Scheffé Tests on differences between means
Grade level

*D	Value D	*S	P<
$E_1 - E_2$	1.4	2.31	
$E_2 - E_3$	2.7	2.31	0.001
$E_3 - E_4$	-1.8		
$E_{1-2} - E_{3-4}$	2.3	1.803	0.001

*See Kerlinger and Pedhazur (1973) p. 129.

stress placed on attitudinal aims, together with the policy of the N.S.W. Schools Board that attitudes should be included in school assessment programmes, has created a major dilemma for N.S.W. science teachers. The results of this study indicate that such a dilemma can be a very real one, particularly for young teachers just out of training.

Out of 23 sets of results obtained from 12 cognitive achievement tests, set by a class of 19 Diploma in Education students, only two produced a coefficient alpha reliability of the order of 0.80 and none had an alpha of 0.85 or above.

In their first year of teaching, these students will be participating in test construction exercises for internal assessment purposes, where their results will be expected to discriminate between individual students. The attitude instrument developed by the class was promising, for an early stage of instrument development, producing an alpha of 0.65, with item analysis indications that it has the potential for further refinement to produce a useful instrument.

However, correlations between the attitude scales and those achievement tests which had reliabilities sufficiently high to allow reasonable interpretation of results, were very low, indicating very little relationship between the attitude as measured by the scale and science achievement as measured by the cognitive tests. Obtaining a set of student results by adding scores from this instrument to results of achievement tests would be of very doubtful validity. In addition, there is the whole complex issue of the unknown degree to which respondents give socially desirable answers, when it is known that the results of such a test will be used for assessment purposes, influencing crucial decisions about their future.

Analysis of the results of the attitude test by grade level showed a predictable and statistically significant upward shift in scores with increasing Grade level, except for grade 8 which had the lowest mean, but the increase in the mean between junior and senior grades was only a moderate magnitude, tentatively suggesting that the influence of five years of high school science is not a major one in developing a belief in the value of conservation of the natural environment.

Further research

The study was a preliminary and exploratory one and wide generalization from the results is not possible. However the results suggested that it is possible for attitude scales for use in school-based evaluation to be developed, but much more research needs to be done into the relationship between attitudes, as measured by such tests, and cognitive achievement, before the validity of their use in pupil assessment can be assured. If further research supports these findings, the validity of adding the results of attitude tests to those of cognitive tests, as indicated in the N.S.W. Schools Board Assessment Guide, is seriously brought into question.

Following on the project reported in this paper, the writer and Dr A.G. Smith of the University of Newcastle have joined forces to carry out a longer term research and development project in school-based testing, in which the development of attitude scales and the exploration of interrelationships between the attributes they measure and their relationship with cognitive achievement, will be a major component.

Over 30 Hunter Valley high schools have agreed to cooperate in the project, which will first provide an analysis service for tests generated in the schools, with assistance to make use of the analyses provided to improve test reliability. An item bank of items of known characteristics will be built up from the tests analyzed for test improvement and instruments in the affective domain will be developed and the relationships between the variables measured by the various tests investigated.

Acknowledgements

The tests used in this study were developed and their results prepared for analysis by the students enrolled in the 1977 Science Curriculum and Method Unit in the Diploma in Education at the University of Newcastle, namely:-

R. Allomes, I. Anderberg, J. Bailey, R. Bates, J. Besoff, L. Brandt, G. Bright, D. Butel, D. Carey, S. Garnham, J. Greenwell, S. Holmes, P. Lund, T. Nguyen, D. Plumb, R. Ramsey, M. Stanwell, D. Tapp

Assistance in the preparation of data for tables in the report was rendered by Mrs N. Weiss.

References

COHEN, J. Some statistical issues in psychological research. In B.B. Wolman (ed.), *Handbook of clinical psychology*. New York: McGraw-Hill, 1965, 102.

KERLINGER, F.N. & PEDHAUZUR, E.J. *Multiple regression in behavioral research*. New York: Holt, Rinehart & Winston, 1973, 128–131.

McBURNEY, S. *The junior secondary science syllabus: An investigation; investigation 4 – attitudes to attitudes.* Directorate of Planning Services, N.S.W. Education Department, 1978.

MADDOCK, M.N. Measuring attitudes held by high school students and minimally educated villagers in Papua New Guinea. An instrument development study. *Research in Science Education,* 1974, 4, 113–129.

NUNNALLY, J.C. *Psychometric theory*. New York: McGraw-Hill, 1967.

SECONDARY SCHOOLS BOARD. *Syllabus in science, years 7–10: Statement of aims and content*. N.S.W. Department of Education, 1973a.

SECONDARY SCHOOLS BOARD. *Syllabus in science, years 7–10: Asssessment guide.* N.S.W. Department of Education, 1973b.

TELFER, R. *Preliminary survey of teaching problems.* Mimeographed paper, Faculty of Education, The University of Newcastle, 1977.

Research in Science Education, 1978, 8, 135-144

LABORATORY PSYCHOMOTOR SKILL DEVELOPMENT USING PHYSICAL AND MENTAL PRACTICE STRATEGIES

W. F. Beasley

Background to the study

Questions have been raised in the past concerning the lack of research into the outcomes of laboratory instruction, particularly those contained within the psychomotor domain (Dietrich, 1976; Doran, 1976; Klopfer, 1971; Ramsey & Howe, 1969). At the tertiary level, Cole (1966) and Nash (1976) have questioned whether the large proportion of University science students' time devoted to laboratory work can be justified. Large sums of money are spent on the instructional process of the laboratory lesson, but only relatively minute sums have been devoted to examining the assumptions underpinning the teaching laboratory.

Reflecting on freshman chemistry, Nash (1976) has questioned the value of laboratory manipulative skills "in an age when the computer interfacing with power instruments may require, in the limit, no more than the insertion of a sample and subsequent collection of a written report on its characteristics" (p. 607). On the other hand, Fritz and Schenk (1974) have maintained that:

> Although a number of purely instrumental methods of analysis have been developed, chemical methods are still very vital and widely used . . . few instrumental methods can match the accuracy and precision of chemical methods (p. 8).

Research activity concerning the development of motor learning has traditionally been the domain of the psychologist, physical educationist, and the industrial educationist. These groups have considered practice to be a key condition in learning motor skills. The conditions of practice may be analysed in terms of at least seven variables identified by Knapp and Dixon (1950). These include:

1. Duration of the practice session.
2. Length of the rest period between practice sessions.
3. The practice method.
4. Speed of movement during practice.
5. Characteristics of the learner.
6. Acttivity of the learner during the time between practice periods.
7. Complexity of the skill.

The amount of practice and subsequent learning has generally been assumed to be a function of the period of time the individual participates, or the number of repetitions that are completed. This emphasis on the overt or physical performance of the task to be learned has been consistent with the philosophy of learning by doing (Oxendine, 1968). However, there now exists sufficient evidence to suggest that physical proficiency might be enhanced by mental practice which precedes, accompanies or follows the performance (Richardson, 1967; Oxendine, 1972).

The term mental practice is used to signify the introspective or covert rehearsal that takes place within the individual (Richardson, 1967; Oxendine, 1968). In the physical performance of a task there is usually some degree of mental activity, while in mental practice neural and muscular responses are evoked. Thus mental practice could be better understood if thought of as passive learning (Singer, 1968).

A considerable body of literature on mental practice and its effects on the acquisition of motor skills has accumulated over the last fifty years. This has been reviewed by Richardson (1967) who suggests mental practice facilitates learning. Richardson postulates that as the learner imagines himself going through the sequence of required actions, there will be minimal innervations of the relevant muscles. The actual kinesthetic feedback which occurs is assumed capable of evoking its associated image so that corrections can be made for the next trial. "Under this construction a more familiar task would have greater gains after mental practice since a clearer model of what constitutes a good performance would have been internalized" (Smyth, 1975, p. 199).

Morrisett (1956) categorized tasks according to the symbolic, perceptual or motor nature of the content and found that predominantly motor tasks were little helped by mental practice. Mental practice may be effective in the early or cognitive stage of skill acquisition (Fitts & Posner, 1967) for a task which has a relatively high symbolical content. Beckow (1967) has suggested that rehearsing a skill mentally may be more effective when the skill requires fine adjustments, control, and precision rather than one which requires a great amount of power for proficient performance.

In the past six years studies have appeared concerning the role of mental practice in the development of vocational psychomotor skills. Schacht (1971) found mental practice had value in acquiring a vocational printing psychomotor skill. It has also been shown that college level males learn as well a manual arc welding task by engaging in a practice session consisting of both mental and physical practice when compared to an equal period of physical practice (Hackler, 1971). This study was significant in that the mental practice technique allowed more efficient use of the workshop facilities.

Since this researcher has found no reported research concerning psychomotor skill development in science education, one is encouraged to draw implications from the areas discussed above for possible instructional techniques. It is possible that such techniques would allow the efficient development and refinement of student laboratory psychomotor skills to a level in which the accuracy of student experimental measurements approaches the precision of the apparatus. Given that success in the science laboratory requires psychomotor skills which may be classified as 'fine', containing a significant symbolic dimension, research seems to imply that:

1. Mental practice when used in conjunction with overt practice may lead to a significant improvement in laboratory performance (Corbin, 1966; Egstrom, 1964; Trussel, 1958).
2. Some experience with a particular psychomotor task is necessary before mental practice can be fully effective (Phipps, 1968; Corbin, 1967).
3. Mental practice can be effectively used with students with intelligence within the range normally found in a public school (Start, 1960).

The purpose of this study was to assess the separate and joint effects of physical and mental practice in developing skill in selected psychomotor tasks.

Statement of hypotheses

General research hypothesis: The effect of physical practice of 'fine' psychomotor skills on chemistry student laboratory performance is moderated by the presence of mental practice.

The six specific hypotheses of interest can be summarized as follows:

There are no significant differences in student laboratory performance when the associated 'fine' psychomotor skills are developed by any of the practice techniques presented in the following matrix (A x B) in Table 1.

Numbers 1–6 in Table 1 indicate the comparisons of practice techniques which were tested in hypotheses 1–6 respectively.

TABLE 1

Comparative practice techniques for the development of laboratory psychomotor skills

	P	*M*	*P + M*
M	1	▨	▨
P + M	2	3	▨
C	4	5	6

Key

P = Physical Practice
M = Mental Practice
P + M = Physical & Mental Practice
C = Control

Procedures

The Chemistry 103 laboratory program at the University of Maryland (Lower Division Chemistry Staff, 1975) emphasizes quantitative volumetric analysis techniques in one quarter of the experiments. This technique involving the use of readily available apparatus requires accurate laboratory measurements for a satisfactory laboratory performance. This performance is dependent upon a high level of student psychomotor skill in carrying out the laboratory tasks of weighing and liquid volume measurement.

Two teaching modules were developed and validated by the researcher to provide instruction in the psychomotor skills involved in the correct use of the analytical balance, the volumetric pipet, and the buret.

The teaching module – Quantitative Volumetric Analysis: Skill Development – was developed for the purpose of providing instruction and physical practice exercises in the correct use of the balance, the pipet, and the buret.

The teaching module – Quantitative Volumetric Analysis: Review and Practice – was developed for the purpose of providing instruction and mental practice exercises in the correct use of the balance, the pipet, and the buret.

The experimental group consisted of those students (approximately 450) who enrolled in two Chemistry 103 lecture sections in the Fall Semester 1977, at the University of Maryland. The Chemistry 103 lecture sessions did not include those students with superior Mathematics SAT scores. These students, of the order of 5–6% of the introductory chemistry student population, formed a separate Chemistry 105 honours course lecture section. The experimental group formed 24 laboratory sections which were randomly assigned to the four treatment groups as follows:

Treatment Group 1 – Physical Practice
Treatment Group 2 – Mental Practice
Treatment Group 3 – Physical and Mental Practice
Treatment Group 4 – Control

The sequence of events during the entire treatment period over three laboratory sessions was as follows:

Week 11. All groups completed the pretest.
Week 12. Groups 1 and 3 completed the physical practice teaching module. Groups 2 and 4 completed a non-related laboratory experiment, Freezing Point Depression (Lower Division Chemistry Staff, 1975).
Week 12–13. In the period of time between the completion of the experiment in week 12 and commencement of week 13, Groups 2 and 3 completed the mental practice teaching module.
Week 13. All groups completed the posttest.

Two criteria, the accuracy and precision of a student's laboratory measurements, were applied in the determination of the proficiency of student performance in the pretest and posttest experiments.

The accuracy and the precision of student performance were determined in the following manner:

1. Each student recorded his/her laboratory measurements on specially prepared data sheets.
2. The student raw data were transferred from the data sheets to computer cards.
3. The composition of the sample chemically analysed by the student was determined by computer analysis on the basis of the student data. This was compared to the known composition of the student sample.
4. The error of each determination was calculated. The relative error expressed in parts per thousand by mass was calculated for each student determination. The accuracy and the precision as defined previously were calculated for both the pretest and posttest laboratory performances.

The calculation procedure described above provided the researcher with two measures of student performance in both the pretest and posttest.

The statistical analyses which were undertaken on the pretest and posttest measures of student laboratory performance are as follows:

1. The reliabilities of the pretest and the posttest measures were established by calculating the correlation coefficient of each trial of student determinations with the other two trial determinations. Each student carried out three trial determinations of the unknown sample in both the pretest and posttest experiments.
2. To measure the change in student performance over the treatment period a multivariate analyses of covariance procedures (with pretest performance as the covariate) was employed. The important assumption of homogeneity of regression in the analyses of covariance was tested by application of the computer program ANCOVA 3 (Dayton, 1973). The analysis of covariance was performed using the computer program MANOVA (Biometric Laboratory, University of Miami).

3. Pairwise comparisons of group means using the Newman Keul method were used to test Hypotheses 1–3.
4. Hypotheses 4–6 were tested by applying Dunnett's test to contrast each treatment group with the control group.

Analyses of student performance

Accuracy and precision criteria

Table 2 summarizes the breakdown of variables and factors entered in the 2 x 2 Multivariate Analysis of Covariance design used to compare the accuracy and the precision measures of student performance respectively.

TABLE 2

Variables and factors entered into the Analysis of Covariance Design –Accuracy and precision measures

3 Variables	2 Factors each at 2 levels
1 Criterion Measure	Factor A: Physical Practice
Posttest Accuracy Score or Posttest Precision Score	(Yes, No)
2 Covariate Measures	Factor B: Mental Practice
Pretest Accuracy Score	
	(Yes, No)
Pretest Precision Score	

Table 3 presents a summary of all data for the variables entered into the analysis.

Both accuracy and precision performance measures were found to satisfy the important homogeneity of regression assumption needed for the analysis of covariance procedures used to test treatment effects.

Comparison of group performance – Accuracy criterion

Table 4 presents the findings of the F tests to determine the significance of physical and mental practice on student accuracy in the performance of chemistry laboratory experiments involving the technique of quantitative volumetric analysis.

The results of the F tests, as reported in Table 4, indicate a significant difference among treatment groups associated with both main effects (Physical and Mental Practice) and also with the interaction effect (Physical Practice x Mental Practice). By reference to the adjusted cell means presented in Table 3, the directionality of each significant effect can be judged. It would appear that the significant differences which have been found relate to the accuracy score of the control group. The greater the 'accuracy score' the less proficient the laboratory performancce was considered to be.

Students in Treatment Groups 1, 2 and 3 established the composition of the unknown

TABLE 3

Summary of cell size, means and standard deviations for variables in the laboratory performance ANCOVA analysis

Factor *A*	 *B*	*Cases*		*Variable* *Pretest accuracy (ppt)*	 *Pretest precision*	 *Posttest accuracy (ppt)*	 *Posttest precision*
1	1	87	M	147.103	43.397	24.279 (23.908)*	13.727 (13.527)*
			sd	120.984	74.934	21.086	12.840
1	2	84	M	113.729	36.489	24.145 (24.313)*	15.862 (15.983)*
			sd	64.650	45.646	21.002	16.481
2	1	93	M	106.888	33.313	28.371 (28.790)*	19.728 (20.033)*
			sd	71.168	44.112	20.789	23.811
2	2	96	M	110.170	41.381	45.042 (44.824)*	40.132 (39.902)*
			sd	76.070	45.768	67.092	70.077

Key: A = Physical Practice 1 = Yes
B = Mental Practice 2 = No

*Adjusted mean on the basis of the two covariates

TABLE 4

Results of F test to determine the effect of physical and mental practice on the accuracy of student laboratory performance

Source	*df*	*MS*	*F*	*P less than*
Within Cells	354	1513.75		
Regression	2	3199.96	8.180	0.004
Physical Practice	1	12048.39	7.959	0.005
Mental Practice	1	8331.91	5.504	0.020
Physical Practice x Mental Practice	1	6880.41	4.545	0.034

sample with a relative error of the order of 2.5% (25 ppt) which compares to a relative error for the control group of the order of 4.5% (45 ppt).

The significant interaction effect (Physical Practice x Mental Practice) suggested that the six specific null hypotheses inherent in the research design could be investigated for differences in group performance on the accuracy measure.

A summary of the results relating to the testing of the six hypotheses inherent in the research design is presented in Table 5.

TABLE 5

A summary of statistics for the testing of hypotheses 1–6 relating to student accuracy measures

Hypotheses	*Treatment*		*Statistical*	*df*	*Significant*
	Group	*Means*	*Test* Value*		*at 0.05 level*
1	26.79, (mp)	24.31 (pp)	0.6039	88	no
2	23.91, (m&pp)	24.31 (pp)	0.0974	88	no
3	23.91, (m&pp)	26.79 (mp)	0.7014	88	no
4	24.31, (pp)	44.82 (c)	–3.628	178	yes
5	26.79, (mp)	44.82 (c)	–3.186	187	yes
6	23.91, (m&pp)	44.82 (c)	–3.631	181	yes

*Hypotheses 1–3 : Newman Keul Test
Hypotheses 4–6 : Dunnett's Test

Key:
mp = Mental Practice Group
pp = Physical Practice Group
m&pp = Mental and Physical Practice Group
c = Control

The next section presents the analysis of covariance findings for the treatment effects on student performance regarding precision of results.

Comparison of group performance – Precision criterion

Table 6 lists the results of F tests to determine the significance of physical and mental practice on student precision in the performance of Chemistry student laboratory experiments involving the technique of quantitative volumetric analysis.

The results of the F Tests, reported in Table 6 indicate a significant difference among treatment groups associated with both main effects (Physical and Mental Practice) and also with the interaction effect (Physical Practice x Mental Practice). By reference to the adjusted cell means presented in Table 3, the directionality of any significant effect can be estimated. It would appear that the significant differences which have been found relate to the precision score of the control group. The greater the precision score the less proficient the laboratory performance is considered to be. The significant interaction effect suggested that the six specific null hypotheses inherent in the research design could be investigated further for differences in group performance on the precision measure.

A summary of the findings relating to the testing of Hypotheses 1–6 is presented in Table 7.

TABLE 6

Results of F test to determine the effect of physical and mental practice on the precision of student laboratory performance

Source	*df*	*MS*	*F*	*P less than*
Within cells	384	1555.27		
Regression	2	2204.68	1.412	0.244
Physical Practice	1	20392.39	13.112	0.001
Mental Practice	1	12009.08	7.722	0.006
Physical Practice x Mental Practice	1	6714.00	4.317	0.038

TABLE 7

A summary of statistics for the testing of hypotheses 1–6 relating to student precision measures

Hypotheses	*Treatment*		*Statistical*	*df*	*Significant*
	Group	*Means*	*Test* Value*		*at 0.05 level*
1	20.03, (mp)	15.98 (pp)	0.9729	88	no
2	13.53, (m&pp)	15.98 (pp)	0.5885	88	no
3	13.53, (m&pp)	20.03 (mp)	1.5614	88	no
4	15.98, (pp)	39.90 (c)	–4.059	178	yes
5	20.03, (mp)	39.90 (c)	–3.462	187	yes
6	13.52, (m&pp)	39.90 (c)	–4.536	181	yes

*Hypotheses 1–3 : Newman Keul Test
Hypothese 4–6 : Dunnett's Test

Key:
mp = Mental Practice Group
pp = Physical Practice Group
m&pp = Mental and Physical Practice Group
c = Control

Summary of findings

Hypotheses 1–3 related to the joint and separate effects of physical and mental practice on student laboratory performance. For both accuracy and precision performance criteria, no significant differences were found among the three treatment groups that experienced either physical practice, mental practice, or physical and mental practice.

Hypotheses 4–6 compared the joint and separate effects of physical practice and mental practice on laboratory performance to control group performance, where the control group experienced no distinct practice phase. For both the accuracy and precision measures of laboratory performance, the empirical evidence indicated a significant difference in performance for all three treatment groups when compared to the control. The directionality in cell means indicated that treatment groups which experienced either physical, mental, or physical and mental practice performed significantly better than the control group on both laboratory accuracy and precision measures.

Some conclusions

This study has demonstrated the feasibility of improving the laboratory performance of freshman chemistry students through planned skill practice activities. However, no differences were found among the specific practice strategies which were employed. Mental practice of laboratory techniques represents the most efficient strategy in terms of the necessary laboratory resources.

Although mental practice alone was not expected to be as effective as physical practice, this study's findings proved to be an exception to the implications of earlier reported literature. However, if one accepts the argument that prior experience with the appropriate apparatus during the pretest experiment and/or high school courses was significant, then this study is consistent with the implications of the proposed model. Thus mental practice, to be fully effective, should be associated with prior experience of the task.

Evaluation of such findings should relate to the degree of improvement which was found relative to the investment of student time and laboratory resources in the practice activities. A more detailed discussion of the study is provided elsewhere (Beasley, 1978).

References

BEASLEY, W.F. The effect of physical and mental practice of psychomotor skills on chemistry student laboratory performance. Unpublished doctoral dissertation, University of Maryland, 1978.

BECKOW, P.A. A comparison of the effectiveness of mental practice upon learning of two gross motor skills. Unpublished masters thesis, University of Oregon, 1967.

COLE, A.R.H. Laboratory courses and laboratory design. In *Conference on the teaching of chemistry*. Sydney: Federation of Australian University Staff Associations, 1966.

CORBIN, C. The effects of mental practice on the development of a unique motor skill. *Proceedings of the National College of Physical Education for Men*, 1966, 69, 100–102.

CORBIN, C. The effects of mental practice on skill development after controlled mental practice. *Research Quarterly*, 1967, 38, 534–538.

DAYTON, C.M. *Analysis of covariance ANCOVA 3.* University of Maryland Computer Center Library, 1973.

DIETRICH, M.C. MLT and MT student performance on a preliminary psychomotor test battery. *American Journal of Medical Technology*, 1976, 42, 461–468.

DORAN, R.L. Implications for measurement and evaluation from the trends of science education. *Science Education*, 1976, 60, 199–209. (b)

EGSTROM, G.H. Effects of an emphasis on conceptualizing techniques during early learning of a gross motor skill. *Research Quarterly*, 1964, 35, 472-481.

FITTS, P.M. & POSNER, M.I. *Human performance*. Belmont, California: Brook-Cole, 1967.

FRITZ, J.S. & SCHENK, G.H. *Quantitative analytical chemistry* (2nd ed.). Boston: Allyn & Bacon, 1974.

HACKLER, C.M. The relative effectiveness of physical and mental practice on the learning and retention of a selected psychomotor task. Unpublished doctoral dissertation, University of Maryland, 1971.

KLOPFER, L. Science evaluation. In B. Bloom (ed.), *Handbook of formative and summative evaluation of student learning*. New York: McGraw Hill, 1971.

KNAPP, C.G. & DIXON, W.R. Learning to juggle: A study to determine the effect of two different distributions of practice on learning. *Research Quarterly*, 1950, 21, 331-336.

LOWER DIVISION CHEMISTRY STAFF, University of Maryland. *Laboratory chemistry (Vol. 1)*. Minneapolis: Burgess, 1975.

MORRISETT, L.N. The role of implicit practice in learning. Unpublished doctoral dissertation, Yale University, 1956.

NASH, L.K. Reflections on freshman chemistry. *Journal of Chemical Education*, 1976, 53, 606-608.

OXENDINE, J.B. *Psychology of motor learning*. Englewood Cliffs, New York: Prentice Hall, 1968.

OXENDINE, J.B. Physical education. In R.B. Singer (ed.), *The psychomotor domain: movement behaviour*. Philadelphia: Lea and Febeger, 1972.

PHIPPS, S.J. Effects of mental practice on the acquisition of motor skills of varying complexity. Unpublished masters thesis, Pennsylvania State University, 1968.

RAMSEY, G.A. & HOWE, R.W. An analysis of research on instructional procedures in secondary school science: part I, outcomes of instruction. *The Science Teacher*, 1969, 36, 62-70.

RICHARDSON, A. Mental practice: A review and discussion, part one. *Research Quarterly*, 1967, 38, 95-107.

SCHACHT, R.C. An experimental investigation to analyse the effect of mental practice on the initial acquisition of a motor skill. Unpublished doctoral dissertation, University of Maryland, 1971.

SINGER, R.N. *Motor learning and human performance: An application to physical education skills*. New York: MacMillan, 1968.

SMYTH, M.M. The role of mental practice in skill acquisition. *Journal of Motor Behaviour*, 1975, 7, 199-206.

START, K.B. The relationship between intelligence and the effect of mental practice on the performance of a motor skill. *Research Quarterly*, 1960, 31, 644-649.

TRUSSEL, E.M. Mental practice as a factor in learning a complex motor skill. Unpublished doctoral dissertation, University of California, Berkeley, 1958.

Research in Science Education, 1978, 8, 145-156

SKILLS AND SKILLS TESTS IN EXPERIMENTAL PHYSICS

Lawrence G. Little and Rob W. de M - Mackay

Introduction

The identification of laboratory and workshop skills relevant to the teaching of high school physics has been a matter of high priority within the Physics Department of Sydney Teachers' College since 1974. This priority has arisen from at least four considerations.

Firstly, there is the long standing and undesirable situation in which less than 10% of university graduates entering the College 'end-on' Diploma of Education possess any physics beyond the first year of their degree course. Of the others, the majority will have taken only the first year Life Sciences Physics course, which may be seen as a relatively 'soft option'. There is concern whether these graduates will be adequately equipped to teach a junior Science Syllabus which acknowledges a basic dependence on the physical sciences.

Secondly, this priority is related to the proliferation of laboratory and demonstration equipment which has occurred in N.S.W. schools since the influx of Commonwealth Grant money. Much of the equipment has been of an electrical or electronic nature and has been developed and acquired at a rate which has outstripped the training of teachers to use and maintain it effectively in the face of quite minor deterioration. The current emphasis on accountability provides an added incentive to produce teachers able to handle the newer solid state devices effectively and with confidence.

The third consideration leading to the identification and listing of skills' objectives relates to the time factor in an 'end-on' Dip.Ed. course. At Sydney Teachers' College, science graduates are provided with an elective programme in which they can choose from nine to forty two hours physics during their one year course. This makes it desirable to provide all students with the opportunity of assessing their position and progress realistically. A statement of objectives also makes it possible for students to avail themselves of open laboratory opportunities in which to overcome individual deficiences. To this end, the authors have prepared and published *Basic Skills in Physics*, which is a manual related to skills seen as essential to the teaching of science up to the level of year 10.

Fourthly, it is hoped to be able to assess all Science students in the College in the psychomotor as well as in the other areas of performance. The statement of objectives relating to basic skills is seen as a prerequisite to developing suitable test items. A major part of this paper deals with this testing.

Although this work is principally concerned with skills relevant to teaching junior science, the authors are currently concerned with extending it to include senior school physics and the use of computers in physics education.

The extension of this work to include the analogue computer began with the acquisition of an EAI 180 Analog Hybrid computer in 1974. This was initially to provide simulation in the area of atomic and nuclear physics where it has become difficult to match theoretical aspects with convenient laboratory experience. It has more recently been used in the Biology Department for simulation of ecological systems and in the first year Physics laboratory to provide further experience with Newtonian mechanics. Programmes for use in undergraduate chemistry, earth science and social science have also been developed. In 1976 a two day course with 'hands-on' experience was given to a Year 10 class at a Sydney High School where it was found that the pupils were able to use the computer with relative ease.

The EAI 180 provides students with a quick and efficient means of model testing, relies heavily upon an understanding of the physical significance of terms in an algebraic equation without needing to solve it and gives the user an answer in graphical form. It follows that skills associated with its use will involve identifiable cognitive and psychomotor elements and that the analog computer will provide a suitable medium to examine student performance. Students' interest in applications such as the Schroedinger wave equation has been encouraging.

Identification of skills and objectives

Sund and Trowbridge (1973) have identified five sets of skills as essential in the knowledge and performance of science. Of these, the concern here is largely with manipulative skills. Showalter (1974) identified 15 instruments which are useful in skilled physical activities. In 1974 the five lecturers of the Physics Department contributed to a list of instrumental and workshop skills considered essential for the teaching of physics in the high school science programme up to year 10.

This list was considerably modified as the result of review and consultation with teachers. The table of contents of the resulting manual, *Basic Skills in Physics*, is shown in Appendix 1. Objectives relevant to each of the instrumental skills are listed in this manual and these are still subject to modification and extension.

Skills tests in experimental physics

The development of these tests goes back to 1963 when a Practical Examination Sub-Committee of the N.S.W. Secondary Schools Board Science Syllabus Committee was investigating ways of measuring experimental science skills, with a view to a state-wide practical examination in Form 4 (Year 10) as part of the Science requirements for the School Certificate. The problems which arose led to the abandonment of the quest and the disbandment of the committee. It so happened that L.G. Little and R.W. Maclay organised a trial under the scrutiny of the Syllabus Committee for nearly 200 Year 9 pupils at Sydney Grammar School in August 1963.

Further details of this trial, similar trials and a later (1968) N.S.W. state-wide trial practical examination were reported by Maclay (1971) in a thesis which put the practical examination into the context of a taxonomy of educational objectives in the psychomotor domain. The instrument used for this examination, and the taxonomy had been described earlier (Maclay, 1969), quoted in Tisher, Power and Endean (1972) and the potential applications and tentative developments in Maclay (1974). Use was also made of the taxonomy in the Australian Science Teacher Education Project (ASTEP) trials and final publication (Northfield, 1976).

Correlations between scores in these earlier trials 1968 and 1973 gave qualitatively similar results to studies in a number of learning areas associated with manual or related skills (e.g. Baldwin, 1971) namely that correlation coefficients between scores in practical and related cognitive tests were low to negative. Critical examination of test items, however, emphasises the difficulty of examination of specifying items as purely psychomotor i.e. free of cognitive (and, indeed, affective) components without making them trivial. The gross correlations of total scores merely confirm that different objectives were being tested in ways which might suggest that scores were being given in test items for which validity and reliability were suspect.

In a wider context of attention to the taxonomy (Macley, 1977) progress seems to have been made in Physics tests recently. All tests described were of the 'stations' kind providing apparatus at each for appropriate attention directed by the test paper carried by the subject. Five minutes were allowed at each station. Communications tests followed in two of the instruments.

1. Prescribing objectives (before the test)

This activity encouraged the student to consider past practical experiences and to refer to any records of these, together with their context of concept, process and other essential content. It can be seen from these that a prescription had been made which encouraged a valid selection of test items. Whether this represents a valid selection from the psychomotor viewpoint is another matter which is considered under 4.

2. Analysing cognitive and psychomotor components

The accompanying list of objectives in Appendix 2 anticipates the development of a scoring scheme which is realistic to the items devised by the teacher and the responses of the learner. Before scoring it was convenient to record the results of all subjects in a given test item so that consistent scores could be allocated in the interests of reliability, all results being available for immediate inspection.

3. Conducting intercorrelational studies

As suggested in the earlier third paragraph, low correlations between psychomotor and cognitive scores have been usually expected. Some intercorrelations for the three tests are presented in Appendices 3 and 4. The first gross observation was the large number of relatively high correlations. Those of psychomotor interest were chosen for attention, noting again that many of the lower correlations were associated with cognitive/psychomotor scores. Further details may be obtained from the authors.

3.1. *Of correlation coefficients* $\geqslant 0.5$

207 1976 (N = 22) 12 stations and a communications test.[1]

1P (4th highest mean). It was required to vary 4 only CRO controls for an input signal, draw the trace and calculate and record the peak voltage and period.

5P (10th highest mean). It was required to weigh an object to 10^{-7} kg.

9P (2nd lowest mean). It was required to connect a circuit to measure AC current through a coil whose inductance and the equation for calculating its reactance were supplied.

12P (4th highest mean). It was required to find the two current readings through the loop which produced balance for two required current settings through the coil in a displayed current balance circuit.

---

[1] The numbers 207 and 107 describe student sections in first and second years.

207 1977 (N = 14) 10 stations and a communications test

4P (3rd highest mean). It was required to identify the potential of a charged electroscope, and determine the temperature 2 cm above a lamp globe using a thermistor probe and a supplied resistance/temperature graph.

6P (4th highest mean). It was required to set a grooved stick at an angle to the horizontal and time the rolling of a ball from five points using a stop watch.

9P (2nd lowest mean). It was required to determine the mean and uncertainty in the tachometer readings of frequency of rotation of a disc.

11P (4th lowest mean). It was required to do some interpretations and calculations of instantaneous linear velocity, acceleration and angular velocity from the results obtained in 6P.

107 1977 (N = 18) 10 stations

6P (6th highest mean). It was required to select components and connect a circuit from the supplied diagram, to be scored by an observer.

9P (8th highest mean). It was required to select apparatus and measure the focal length of a lens.

10P (4th highest mean). It was required to measure the wave length and amplitude of an AC input to a CRO at specified settings of voltage and time base.

3.2 *Those of correlation coefficient < 0.1*

207 1976 (N= 22) 10P (highest mean). It was required to polaroid photograph two horizontal runs of a trolley each accelerated by different forces, the timer being a fixed "blink"-rate.

207 1977 (N = 14) 2P (lowest mean). It was required to measure air pressure above atmospheric at a fixed point on a Venturi tube for a required setting of the fan.

7P (2nd highest mean). It was required to use a buzzer timer to produce a tape carbon record of the motion of the circumference of a wheel rotating about a horizontal axis.

3.3 *The results*

Detailed study of scores confirmed as the results suggested that the test items, for various reasons, scatered the subjects well. The intercorrelations between items and of items with the various totals gave further hints on those which might be modified. The subjects of tests were given the opportunity of comment on the items as to their confidence with them, the sufficiency of time allowed and any administrative difficulties which might have affected results. Beyond the indication of insufficient time for an occasional item by a small majority of subjects there was general approval of each test. The effect in no case prevented many from getting reasonable results. Further details may be obtained from the authors.

4. Validity of the test items. Reliability of scoring

An example of a grid of distribution across the psychomotor domain taxonomy is given in Appendix 5. This and a comparison of the test items with the objectives was a guide to face validity. The 107 1977 test was administered to the initial intake of students who would have had varied experiences at different schools. The other two tests represented valid selections from experimental experience. The intercorrelational studies gave a guide to content validity of the components of each test item. How to act on the evidence obtained is complicated by the individuality of the items and their components, hence further work on this is proceeding for revision of the instruments. The small number of test items and large variance gave very high values of reliability as determined by the use of right-wrong item statistics with Kuder-Richardson formulae 20/21 (Roscoe, 1969). Other kinds of reliability measures (e.g. test-retest, equivalent forms, split-half) were not applicable because the time for comparable tests was not available and the halving of the instruments was not appropriate, each item representing a unique experience. The intercorrelations for the split-half as between the psychomotor and cognitive components of each test gave 0.72 for 207 (1976), 0.71 for 207 (1977) and 0.73 for 107 (1977), with corresponding reliabilities of 0.83, 0.83 and 0.84 respectively.

5. Scoring

Only occasional use was made of observers. Most of the data was objectively scored from the paper, having regard for the teacher's anticipation of problems, trying the final items just before the subjects did and restoring equivalent conditions for each subject if an unexpected crisis occurred. The subjects were reminded of the importance of the final instruction at the end of each item which usually restored equivalent conditions. This procedure avoided the labour-intensive problem of watching each site too closely.

Conclusion and acknowledgments

The authors recognise the developmental state of this testing and would welcome comment on the progress to date and the kinds of statistical measures for such test items. The help of Mr R.S. Horsfield (Cumberland C.A.E.) in arranging computer facilities and cooperating in an earlier test at his College assisting the development of 207 (1976) is appreciated. Likewise, the cooperation of Mr R.G. Davis and Mr I.J. Cooper of Sydney Teachers College Physics staff is appreciated.

The whole area of providing adequate experience for science teachers in handling experimental equipment and the testing for attainment of skills is of value to their developing confidence. Their ability to transfer confidence and necessary skills to those who learn from them may be expected to follow. Involvement in experimental activity of challenge and the need to face testing of skills have produced good responses from learners at all levels, both at school and at tertiary institutions. The authors hope that teachers at all levels will try the motivational impact of these practices.

References

BALDWIN, T.S. Evaluation in learning in indudstraial education. In B.S. Bloom *et al.*, *Handbook on formative and summative evaluation of student learning.* New York: McGraw-Hill, 1971, 855–906.

MACLAY, R.W. deM. A taxonomy of educational objectives in the psychomotor domain. *Educational measurement and assessment.* Melbourne: Australian College of Education, 1969, 127–148.

MACLAY, R.W. deM. The practical examination as a test of science skills with particular reference to secondary courses in N.S.W. and especially to the School Certificate Science Course. Unpublished thesis, University of N.S.W., 1971.

MACLAY, R.W. deM. Recent developments in psychomotor domain taxonomies and their application to experimental science. *Research in Science Education*, 1974, 4, 77–80.

MACLAY, R.W. deM. Taxonomies of educational objectives in the psychomotor domain. *Rescent*, 1977, 1(1), 12–25.

NORTHFIELD, J. (ed.). *ASTEP – A project in teacher education.* Monash University, 1976, L. 3/3 and 9.

ROSCOE, J.T. *Fundamental research statistics for the behavioural sciences.* New York: Holt, Rinehart & Winston Inc., 1969, 101–108.

SHOWALTER, V. Program objectives and scientific literacy. *Prism II*, 1974, 2(4), 7.

SUND, R.B. & TROWBRIDGE, L.W. *Teaching science by inquiry in the secondary school.* Columbus: Charles E. Merrill, 1973, 188–190.

TISHER, R.P., POWER, C.N. & ENDEAN, L. *Fundamental issues in science education.* Sydney: John Wiley, 1977, 148–162.

Appendix 1

TABLE OF CONTENTS OF MANUAL

BASIC SKILLS IN PHYSICS

CONTENTS

Introduction

Unit 1 Ray box 1

Unit 2 Photographic Methods 4

Unit 3 Ripple Tank 8

Unit 4 The Microscope 12

Unit 5 The Telescope 19

Unit 6 Spectroscope 27

Unit 7 Spectrometer 29

Unit 8 Young's Experiment 34

Unit 9 Water Equivalent 36

Unit 10 Timers 39

Unit 11 Multimeter and Troubleshooting 44

Unit 12 Soldering and Construction 47

Unit 13 Cathode Ray Oscilloscope 49

Unit 14 Plotting Fields 53

Unit 15 Sound Waves 56

Unit 16 Graphical Methods 58

Unit 17 Radio Reception 63

Unit 18 Static Electricity 66

Appendices:

Appendix 1 Science Skills and the Psychomotor Domain

Appendix 2 The Blinkey

Appendix 3 S.I. Units

Appendix 4 Spectral Data

Appendix 5 Safeguards in the Science Laboratory.

Appendix 2

OBJECTIVES FOR PSYCHOMOTOR SKILLS TEST

SYDNEY TEACHERS COLLEGE

PHYSICS DEPARTMENT
PSYCHOMOTOR SKILLS TEST 1
(including ability to follow instructions)
(Dip Teach. Mar. 1977)

OBJECTIVES (P = psychomotor component, C = cognitive component)

1. P. Ability to use callipers to measure the diameter of a sphere in SI units several times in 3 minutes. Ability to read vernier.
   C. Ability to calculate mean diameter from more than one reading and to report it to figures of significance.

2. P. Ability to draw an ellipse with apparatus and method provided. Ability to measure estimate of major axis.
   C. Knowledge of meaning of major axis.

3. P. Ability to use a balance to weigh accurately to 1 gram.
   C. Ability to record result in kilograms.

4. P. Ability to mix without spilling and to stir till all in solution.
   P/C Ability to read thermometer record of minimum temperature reached.
   P/C Ability to estimate between graduations and to record result.

5. P. Ability to charge a rod and to bring it near a charged electroscope.
   C. Ability to identify the charge on an electroscope as the same or different. Ability to describe the reason for this.

6. P. Ability to connect a circuit from diagram, including selection of identified components.
   C. Ability to identify the + and – terminals of a cell and interpret component symbols.

7. P. Ability to draw a circuit with labelled components, using a connected circuit.
   C. Ability to associate standard symbols and conventions with circuit components.

8. P. Ability to compare approximately the current induced in two coils with the aid of a bar magnet, using similar conditions of induction.
   C. Ability to comment on likely causes of observational effects resulting from different wiring.

9. P. Ability to measure the focal length of a converging lens by any method.
   C. Ability to describe the method used and record the result.

10. P. Ability to adjust the CRO volts/cm and time/cm settings to prescribed values.
    C. Ability to interpret the resulting AC trace in terms of amplitude and wave length.

Appendix 3

207 1976 CORRELATIONS

(< 0.1 underlined, > 0.5 *, comments in text.)

Question	*P prec*	*P comm*	*P total*	*C prec*	*C comm*	*C total*	*Total*
1P	0.712*	0.196	0.646*	0.431	0.229	0.413	0.600*
1C	0.419	0.011	0.363	0.494	-0.064	0.303	0.383
2P	0.297	0.498*	0.600*	0.276	0.267	0.333	0.289
2C	0.267	0.071	0.201	0.298	0.172	0.283	0.289
3P	0.368	0.136	0.362	0.096	0.170	0.134	0.281
3C	0.376	0.221	0.464	0.644	0.143	0.646*	0.603*
4P	0.313	-0.015	0.230	-0.058	0.213	0.082	0.128
4C	0.274	0.105	0.221	0.096	0.364	0.273	0.230
5P	0.647*	0.091	0.567*	0.593*	0.303	0.564*	0.621*
5C	0.603*	-0.228	0.379	0.602*	0.408	0.512*	0.495*
6P	0.460	0.059	0.323	0.410	-0.090	0.265	0.370
6C	0.186	-0.045	0.030	0.193	-0.206	0.076	0.090
7P	0.525*	0.015	0.436	0.510*	-0.083	0.314	0.468
7C	0.225	-0.171	0.205	0.645*	0.000	0.382	0.306
8P	0.416	0.095	0.325	0.170	-0.001	0.142	0.271
8C	0.093	-0.061	0.098	0.406	0.277	0.411	0.224
9P	0.547*	0.336	0.567*	0.382	0.496*	0.624*	0.587*
9C	0.509*	0.098	0.433	0.512*	0.318	0.646*	0.618*
10P	-0.094	0.021	-0.052	0.139	0.000	0.092	0.020
10C	-0.130	0.196	0.130	0.151	0.423	0.288	0.148
11P	0.581*	-0.214	0.364	0.311	0.268	0.343	0.367
11C	0.391	0.039	0.343	0.648*	0.085	0.540*	0.478
12P	0.437	0.406	0.514*	0.424	0.093	0.390	0.476
12C	0.282	0.232	0.317	0.204	0.141	0.223	0.337
P prec	1.000	0.241	0.877*	0.683*	0.368	0.672*	0.867*
C prec	0.683*	0.148	0.687*	1.000	0.421	0.920*	0.846*
P comm	0.241	1.000	0.585*	0.148	0.203	0.233	0.477
C comm	0.368	0.203	0.497*	0.421	1.000	0.726*	0.598*
Prac & comm P	0.877*	0.585*	1.000	0.687*	0.497*	0.719*	0.946*
Prac & comm P	0.672*	0.234	0.719*	0.920*	0.726*	1.000	0.885*
Total	0.867*	0.477	0.946*	0.846*	0.598*	0.885*	1.000

Key: 1P = Item 1 psychomotor component
1C = Item 1 cognitive component
comm = communications test
prec = practical test.

Appendix 3 – contd.

207 1977 CORRELATIONS

(< 0.1 underlined, > 0.5*, comments in text.)

Question	23 *P tot*	24 *C tot*	25 *Total*
1P	0.453	0.400	0.448
1C	0.366	0.693*	0.590*
2P	0.126	0.067	0.113
2C	0.285	0.186	0.246
3P	0.319	0.085	0.206
10C	0.331	0.691*	0.651*
4P	0.785*	0.313	0.604*
4C	0.804*	0.521*	0.810*
5P	0.494*	0.072	0.321
5C	-0.016	0.439	0.214
6P	0.751*	0.487	0.682*
6C	0.757*	0.601*	0.743*
7P	0.110	0.132	0.123
7C	0.102	0.290	0.198
8P	0.465	-0.015	0.254
13C	-0.105	0.185	0.036
9P	0.558*	0.764*	0.709
9C	0.637*	0.616*	0.673*
11P	0.674*	0.709*	0.746*
11C	0.651*	0.696*	0.719*
12P	0.427	0.202	0.347
12C	0.248	0.610*	0.404
P tot	1.000	0.714*	0.932*
C tot	0.714*	1.000	0.918*
TOTAL	0.932*	0.918*	1.000

Appendix 4

107 CORRELATIONS

(> 0.5*, comments in text.)

Question	P_{total}	C_{total}	*Total*
1P	0.388	0.386	0.419
1C	0.398	0.626*	0.573*
2P	0.467	0.148	0.313
2C	0.568*	0.375	0.498
3P	0.363	0.200	0.294
3C	0.653*	0.638*	0.693*
4P	0.288	0.433	0.396
4C	0.421	0.551*	0.529*
5P	0.239	0.166	0.214
5C	0.175	0.360	0.297
6P	0.772*	0.413	0.619*
7C	0.341	0.614*	0.527*
8P	0.554*	0.495	0.565*
8C	0.684*	0.832*	0.822*
9P	0.504*	0.646*	0.628*
9C	0.324	0.707*	0.574*
10P	0.733*	0.494*	0.647*
10C	0.445	0.343	0.419
P tot	1.000	0.728*	0.915*
C tot	0.728*	1.000	0.943*
TOTAL	0.915*	0.943*	1.000

Appendix 5

PSYCHOMOTOR DOMAIN TAXONOMY – DISTRIBUTION OF ITEMS (207 1977)

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1.0 PERCEPTION	1.1 *Sensory stimulation*																				
	1.11 Hearing						✓														
	1.12 Sight	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓										
	1.13 Touch	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓										
	1.14 Taste																				
	1.15 Smell																				
	1.16 Nerve-muscle	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓										
	1.2 *Static cue discrimination*																				
	1.21 Hearing																				
	1.22 Sight	(✓)	✓	(✓)	✓	✓	✓	✓	✓	✓	✓										
	1.23 Touch	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓										
	1.24 Taste																				
	1.25 Smell																				
	1.26 Nerve-muscle	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓										
	1.3 *Dynamic cue discrimination*																				
	1.31 Hearing						(✓)														
	1.32 Sight		(✓)		(✓)	(✓)	(✓)		(✓)	(✓)											
	1.33 Touch					(✓)	✓	✓	✓												
	1.34 Taste																				
	1.35 Smell																				
	1.36 Nerve-muscle					(✓)	(✓)	(✓)	(✓)												
	1.4 *Translation*																				
	1.41 Follow instructions	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓									
	1.42 Record experiences	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓									
	1.43 Response to concept	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓								
	1.44 Isolate concept				(✓)					(✓)	(✓)	(✓)	(✓)	(✓)							
2.0 READINESS	2.1 Mental readiness	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓							
	2.2 Physical readiness																				
	2.3 Emotional readiness																				
3.0 GUIDED RESPONSE																					
	3.1 Imitation																				
	3.2 Trial and error					(✓)			(✓)												
	3.3 Guidance to unknown experience																				
	3.4 Obtaining results, calculations		(✓)							(✓)	(✓)	✓	✓	✓							
4.0 MECHANISM	4.1 Choice of materials				(✓)																
	4.2 Planning																				
	4.3 Exercising skill																				
	4.4 Execution of task																				
	4.5 Interpolation/extrap.																				
5.0 COMPLEX RESPONSE with guidance																					
6.0 COMPLEX RESPONSE without guidance																					

Research in Science Education, 1978, 8, 157-166

QUESTIONING QUESTIONING PART I

Miles A. Nelson

Introduction

One teaching skill mentioned in every science methods book is questioning. Critical thinking is presumably triggered in students by using questions. In view of this importance it is surprising to find little experimental research on the learning effects of teachers questions. Since the 1960's many studies have described the questions used in a science classroom, but few link a teacher's questions to pupil learning (Rosenshine, 1974). This study is an experimental effort to change a teacher's questioning techniques in order to observe the effects on children's learning of science.

Each type of teaching situation has a purpose and it is logical to assume that the types of questions asked should vary with the purpose. For example, a preassessment situation should differ in the types of questions asked as compared to a review situation. The questions asked in a value clarification exercise should differ from those asked in a situation where a concept is being evolved from data. Most research on questioning does not control the instructional situation in which questions are asked (Beisenherz & Tucker, 1974). This research focuses on the class discussions which take place after a laboratory activity; hereafter called a postlaboratory discussion.

Most of the newer primary school science curricula stress the importance of children acquiring science processes, such as observing, inferring and verifying, as well as acquiring science content. The science processes are skills, and research stresses the importance of practice and reinforcement as mechanisms for learning skills. Applying these ideas to the science classroom, postlaboratory discussions are excellent situations in which observations, inferences, and verifications can be elicited using questions. Verbal interactions can keep the learner active, provide practice, and reinforcement in the use of such skills. Unfortunately, many primary school science teachers do not have postlaboratory discussions, or if discussions are held, it is the teacher who does most of the observing, inferring, and verifying. If this is the case, one of the major goals of the newer primary science curricula will not be achieved.

The problem

What effects do postlaboratory discussion tactics have on sixth grade children's learning of science content. The study reported here is not yet complete and is part of a larger project.

Theoretical framework

One criticism of past science education research is that it has no theoretical foundation (Medley & Mitzel, 1963). This research was specifically designed to test a theory generated from prior research. Space does not permit a full elaboration of the theory (Nelson, 1978). Some of the more important elements will be described.

Any theory has a set of phenomenon which it attempts to describe, explain, and predict. At present the theory is an attempt to describe, explain, and predict children's learning of science concepts and processes which have been presented through a laboratory experience. By a laboratory experience is meant each learner directly encounters the phenomenon to be learned, as opposed to reading about the phenomenon, or opposed to watching the teacher or another learner experience the phenomenon. Learning is defined as a change in behaviour from which it can be inferred that knowledge has been acquired and can be applied.

A theory is also capable of generating testable propositions. One dictionary defines propositions as a formal statement of either a truth, or an operation to be performed. It is in this latter sense that the word proposition is used. The theory enables a proposition to be generated which describes how a postlaboratory discussion should be conducted to achieve a given instructional objective. In this research a testable proposition was developed for these instructional objectives:

> When asked, children should be able to state the attributes of (a concept)
>
> Given a series of unknowns, children should be able to identify examples of (the concept) and state reasons for their choices.

The testable proposition is depicted in Figure 1. Its interpretation is as follows. Each block is called a tactic and depicts the verbal interaction, questions and reactions, between four or more people in the classroom. The topic of the discussion is depicted beneath the tactic.

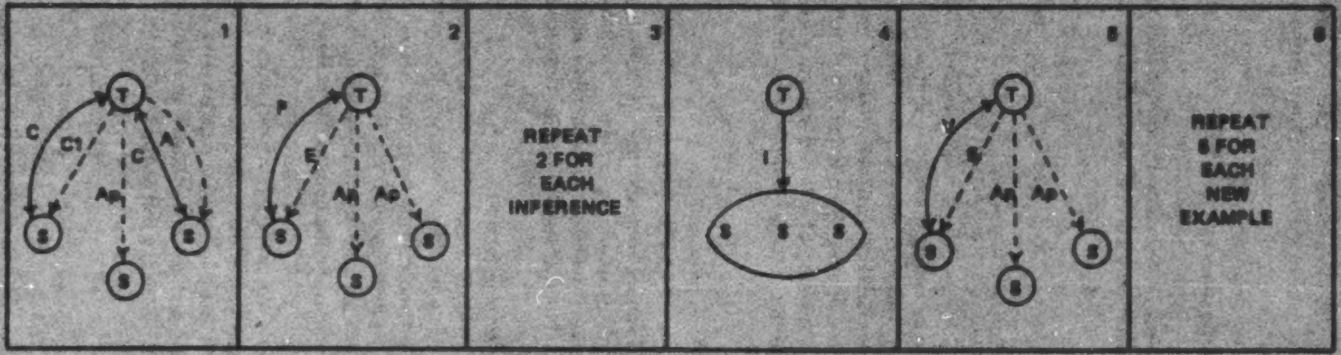
The first step in the postlaboratory discussion is to gather observations from the students and make them visible. This is done by writing their observations on the board. In this tactic many students are asked for their observations. In doing so, the students are practising the skill of observation. The teachers reactions to student responses are Accepting and Calling for Clarification. A Rejecting reaction would be used only if the students gave inferences instead of observations. Writing observations on the board helps to make the transition from the concrete to the abstract and makes possible the next step in the postlaboratory discussion that of noticing regularities in the data.

The second tactic in the postlaboratory discussion is to abstract from observations a science concept or principle. In this tactic the teacher is providing practice with inferring. Students offer inferences to explain similarities in the data. Instead of accepting any answer, the teacher reacts by Calling for Evidence, and by Asking Another to Evaluate the inference. These reactions keep students mentally active, and gives students practice in inferring, and supporting inferences with evidence. Having the data visible, that is having written it on the board, makes it easier for students to process the information. A teacher's reactions to student responses are important. Simply accepting or rejecting the actions are inappropriate. More appropriate reactions would be ones which keep the students who responded mentally active.

After several inferences have been offered and examined for credability, the next tactic would be to summarize the proceedings, and to focus on the publicly accepted concept. Summarizing helps to make available the information for the next stage.

The fourth and final tactic is the application of the science concept, to identify unknown phenomenon. To provide practice using the concept new examples are presented and the students must identify the example and give reasons for their identification. In the Science Curriculum Improvement Study (SCIS, 1974) one science concept presented is ecosystem. The teacher either physically or verbally presents new examples of ecosystems and students are then required to tell why it was or was not an ecosystem. This tactic also provides practice in verification and towards testing a new idea in different situations. Placing these tactics in sequential order results in a proposition for a postlaboratory discussion, in which the teacher's objective is to state the concept and its attributes, and to identify unknowns as

Objective: When asked, children should be able to state the attributes of (a concept). Given a series of unknowns, children should be able to identify examples of (the concept) and state reasons for their choices.


1
T
S
S
S
C
C1
C
A
Ap
2
T
S
S
S
P
E
Ap
Ap
3
REPEAT 2 FOR EACH INFERENCE
4
T
I
S S S
5
T
S
S
S
V
E
Ap
Ap
6
REPEAT 5 FOR EACH NEW EXAMPLE

Focus:

1	2	4	5
Collecting Data and making it visible	Identifying Relationship in Data	Providing a label for phenomena	Identifying unknowns as examples of concept

Key: I Informing

Focusing Questions
C Data Collecting
P Data Processing
V Application/Verification

Reactions
A Accepting
C1 Calling for Clarification
E Calling for Evidence
Ap Calling on Another to react

FIGURE 1

Postlaboratory discussion strategy

examples or nonexamples of the concept. Implementation of this discussion strategy should enable the teacher to reach science content goals and to provide practice in the science processes of observing, inferring, and verifying.

Hypothesis

There is a positive relationship between a class' mean gain in learning selected science concepts and the teacher's use of the theoretically derived postlaboratory discussion strategy.

Method

The research plan was to randomly select teachers to instruct their classes using two prepared science units from the Science Curriculum Improvement Study (SCIS) – Ecosystems and Models. Initially one half of the sample taught the Models unit while the remaining half taught the Ecosystem Unit. Subsequently teachers were trained to conduct postlaboratory discussions in accord with the theoretically derived discussion strategy. The teachers returned to their classes and taught the second science unit. Before and after each science unit, each class was sampled to determine the class learning of selected science concepts. Both before and after training in postlaboratory discussion strategies samples of each teacher's postlaboratory discussions were obtained.

Sample

The population to which generalizations are being made is all sixth grade classes in the Perth metropolitan area. A random sample of seven government schools and six independent schools were selected. Thirteen teachers were involved; five taught two classes each, while the remaining eight each taught one class. Five hundred and fifty students were involved.

Instrumentation

The SCIS evaluation supplements contain tests measuring the concepts taught in the program. These were constituted into 2 one hour tests. Each class was randomly divided into three groups of about 10 students each. Each group took the SCIS tests once, either as a pretest, a midtest, or posttest. Scores from each administration were used to estimate a class' average score. The validity of the tests was estimated by a factor analysis. The reliability was estimated using Cronbach's alpha.

For each teacher a tape recording was made of six postlaboratory discussions. These were typed and then analyzed by trained judges using the Classroom Observational Record (COR) (Reynolds *et al.*, 1971). The COR analysis provides data on who does the talking, the types of questions asked, their frequency, and sequence. The reliability of the COR data was determined using the procedure suggested by Medley and Mitzel (1963).

The teaching programme

The Science Curriculum Improvement Study (SCIS) sixth grade units *Models* and

Ecosystems were used. During the first school term 1977, five 4 hour Saturday workshops were used to teach the teachers the major concepts in the two SCIS units and to develop the SCIS philosophy. Each teacher said the concepts were completely new to him.

During the second school term, each teacher was given the SCIS teacher's guide and a kit of prepared SCIS materials. Each taught at least two science lessons per week averaging 50 minutes per lesson. The teachers were given all of the supplementary materials suggested in the SCIS teachers guides. These included student worksheets and minitests.

In between the second and third term teachers were introduced to a minicourse on Postlaboratory Discussions. The minicourse is a self-instruction package which presents the model discussion strategy along with clues on how to implement it. Exercises are also given which require the teacher to audio tape part of the postlaboratory discussion and analyze it to determine its fidelity to the model presented.

During the third term, each teacher was to utilize the minicourse materials and teach another SCIS unit. Thus each teacher taught both SCIS units, one with their natural questioning techniques and one with the theoretically derived discussion tactics.

Data analysis

The research is based on a classes within treatments design. The data are to be analyzed in several ways. First, using the posttest class mean score adjusted for the pretest class mean and amount of time spent on science instruction as dependent variable a residual can be computed. The residual is to be analyzed by ANOVA using Training in questioning as a factor. In addition the residuals are to be used as dependent variable in a multiple regression analysis. The independent variables will be the frequency and types of questions and tactics used during the postlaboratory discussions.

The second analysis recognizes each test to be comprised of concepts. Research has shown that concepts can be acquired at several levels. The level and number of concepts learned by more than half the class can be analyzed in a contingency table using as categories levels of concept learning and training in questioning. The number of concepts learned by more than half the class can be used as a dependent variable in a multiple regression analysis. The dependent variables will be the frequency and types of questions and tactics used during postlaboratory discussions.

Results and discussion

The postlaboratory discussions analyses are not yet complete. However the analyses of test scores and number of concepts learned are presented in Tables 1 to 4. From the data in Table 1 it can be inferred that the factor training in postlaboratory discussions is not significantly influencing the results obtained for the Ecosystems unit. However, it does explain differences in the Models unit, all favoring the group having training in postlaboratory discussions. It should also be noted that this factor accounts for only a small percentage of the variance in class means. This is not surprising because the number of times each teacher practiced the postlaboratory discussion strategies varied. Some did not practice at all while others practiced 9 or more times.

TABLE 1

Probability of significant differences between treatment groups using class means on subtest scores

	Ecosystems			*Models*		
Oneway ANOVA	*(1, 16 df)*	*P*			*P*	
Pretest	Subtest 1	n.s.		Subtest 1	0.02	-T
	Subtest 2	0.001	T	Subtest 2	n.s.	
	Subtest 3	n.s.		Subtest 3	n.s.	
	Total score	n.s.		Total score	n.s.	
Posttest	Subtest 1	n.s.		Subtest 1	0.05	T
	Subtest 2	n.s.		Subtest 2	0.02	T
	Subtest 3	n.s.		Subtest 3	n.s.	
	Total score	n.s.		Total score	n.s.	
Difference scores	Subtest 1	n.s.		Subtest 1	0.008	T
	Subtest 2	n.s.		Subtest 2	n.s.	
	Subtest 3	n.s.		Subtest 3	0.02	T
	Total score	n.s.		Total score	0.09	T
Residuals	Subtest 1	n.s.		Subtest 1	n.s.	
	Subtest 2	n.s.		Subtest 2	0.02	T
	Subtest 3	n.s.		Subtest 3	n.s.	
	Total score	n.s.		Total score	n.s.	
Oneway ANCOVA adjusting for pretest score and amount of time spent on science						
Posttest	Subtest 1	n.s.		Subtest 1	n.s.	
	Subtest 2	n.s.		Subtest 2	0.02	T
	Subtest 3	n.s.		Subtest 3	n.s.	
	Total score	n.s.		Total score	0.03	T

TABLE 2

Number of classes before and after instruction in which 80% or more of sampled group has learned (+) or not learned (–) ecosystem concept

Concept		*No training in discussion*			*Training in discussion*	
		After			After	
		–	+		–	+
Construct food chains (2 items)	Before +	1	0	Before +	0	0
	–	1	7	–	7	2
		$p < 0.08$				
Identify elements of a community (10 items)	Before +	0	0	Before +	0	0
	–	5	4	–	6	3
Identify environmental factors (5 items)	Before +	0	0	Before +	0	0
	–	7	2	–	6	3
Predict the outcome of an experiment using concept evaporation and condensation (2 items)	Before +	0	4	Before +	0	9
	–	0	5	–	0	0
		$p < 0.03$				
Predict the outcome of an experiment using concept water cycle (3 items)	Before +	0	0	Before +	2	0
	–	2	7	–	3	4
		$p < 0.008$				
Predict the outcome of an experiment using concept 0_2–$C0_2$ cycle (4 items)	Before +	0	0	Before +	0	0
	–	4	5	–	4	5
		$p < 0.03$			$p < 0.03$	

TABLE 3

Number of classes before and after instruction in which 80% or more of sampled group has learned (+) or not learned (–) models concept

Concept			*No training in discussion*				*Training in discussion*	
			After				After	
			-	+			-	+
Identify open and closed electric circuits (7 items)	Before	+	1	2	Before	+	1	2
		-	1	5		-	1	5
			p <0.18				p <0.18	
Given a model, make predictions using it (4 items)	Before	+	0	0	Before	+	0	0
		-	4	5		-	4	5
			p <0.03				p <0.03	
Given a magnetic field identify magnet's location (3 items)	Before	+	0	0	Before	+	0	1
		-	8	1		-	7	1
Identify energy sources and receivers in electric circuits (7 items)	Before	+	0	1	Before	+	0	0
		-	8	0		-	7	2
Construct a model to explain given observations (1 item)	Before	+	0	0	Before	+	0	1
		-	4	5		-	1	7
			p <0.03				p <0.008	

TABLE 4

Effectiveness of teacher as measured by number of concepts learned by more than 50% of class

Models Unit (5 concepts)

	Class																	
	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r
hasn't concept	5	4	3	4	2	2	3	0	3	1	3	1	3	3	2	2	2	4
has concept	0	1	2	1	3	3	2	5	2	4	2	4	2	2	3	3	3	1

$\chi^2 = 21.08 \; p < 0.21$

Ecosystems Unit (6 concepts)

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r
hasn't concept	1	3	3	5	5	4	4	1	4	4	4	0	0	5	2	2	1	3
has concept	5	3	3	1	1	2	2	5	2	2	2	6	6	1	4	4	5	3

$\chi^2 = 32.43 \; p < 0.02$

From the data in Table 2 there are significant differences in the number of classes where 50% or more of the sampled group has learned the Ecosystems concepts when comparing the classes whose teachers have had training in postlaboratory discussions to the classes whose teachers have not. The classes without training have acquired more of the concepts. From the data in Table 3 no trends can be observed.

Data in Table 4 clearly shows the effectiveness of teachers as measured by the number of concepts learned by more than 50% of the sampled group. For example, in Class h and Class l, more than 50% of the sampled groups learned 10 of the 11 concepts presented during the experiment, while in Class d, 50% of the sampled group learned only 2 of the 11 concepts.

What is clear from the data in Tables 2–4 is that there are various levels of teaching effectiveness. When the data from the tape recorded postlaboratory discussion is completely analyzed, it is hoped these differences can be explained by the type and frequency of the teacher's questions, and by the frequency and sequence of the discussion tactics.

Refer[illegible]s

BEISENHERZ, P.C. & TUCKER, J. Limitations of research in science teacher questioning behaviors. *Science Education*, 1974, 58, 423–429.

MEDLEY, D. & MITZEL, H. Measuring classroom behavior by systematic observation. In Gage (ed.), *Handbook of research and teaching*. Chicago: Rand McNally, 1963.

NELSON, M. *Questioning questioning: A final report.* 1978 (in preparation).

REYNOLDS, W., ABRAHAM, E. & NELSON, M. The classroom observational record. Paper presented at Annual Meeting of American Educational Research Association. New York, 1971.

ROSENSHINE, B. Teacher competency research. Paper presented at Annual Meeting of American Educational Research Association. Chicago, 1974.

SCIENCE CURRICULUM IMPROVEMENT STUDY. *Ecosystems* and *Models.* Chicago: Rand McNally, 1970.

Research in Science Education, 1978, 8, 167-174

PRIMARY SCHOOL PUPIL'S ABILITY TO SEE INVESTIGABLE SCIENTIFIC PROBLEMS IN EVERYDAY PHENOMENA: THE TEACHER'S ROLE

David J. Symington

Introduction

A previous paper (Symington, 1977) contained a review of the claims that primary school science should be based upon the investigation of problems that pupils discern in everyday phenomena. The paper also reported data which suggested that there are differences in the number of investigable problems proposed by the pupils depending upon the phenomena considered and the context of the observations.

This paper reports on an investigation built upon the earlier work. The study was designed to provide further data on the impact of a period of free play with materials, and the nature of the materials provided, on the pupils' ability to suggest investigable problems. Further, it generated data on the effect of a teacher led discussion following the period of free play on the same pupil ability.

Before describing the details of the study itself, a number of issues which influenced its design will be discussed.

Individual or corporate problems

An important decision which had to be made for the study was whether to focus on the individual producing problems or on the problems being produced. The latter alternative was the one chosen. This was because, firstly, the courses do not suggest that pupils work in isolation - rather group work is encouraged. Further, the classroom climate proposed in these courses is one which encourages communication between pupils. Where such a climate exists ideas can spread so that the problem posed by one pupil can become a starting point for enquiry by a number of pupils, even the whole class.

In this study therefore, the principal focus is not on the individual and his ability to suggest investigable scientific problems. Rather it is on the influence of free play with materials and subsequent discussion on the ability of the pupils, as a class, to see problems from which investigations can begin.

Materials for free play

It was implied in the earlier paper (Symington, 1977) that the choice of materials for pupil free play was considered to be significant. It was hypothesized that, where the materials provided maximize the opportunities pupils have to make contrasting observations, there would be increased potential for investigable problems to emerge. For example, it was believed that pupils provided with a fully grown, ripe lemon are less likely to wonder about whether the thickness of the skin changes during the ripening of lemons than pupils provided with both a ripe lemon and an unripened one.

It was decided therefore that the design for the investigation should provide for a further evaluation of the relative impact of single and contrasting materials.

Class discussion

There were numerous alternatives which could be considered in assessing the influence of class discussions on the ability being considered. The work of Barnes (1976) has focussed attention on the various forms which classroom discussions exhibit. Unfortunately there is not room to fully discuss these issues here. All that will be done is to describe the decisions made.

Firstly, it was decided that the discussion would be conducted and controlled by the teacher. Secondly, that the only moves the teacher would make were soliciting response (s.r.), exercising control (e.c.), asking clarifying questions (c.q.), summarizing the various views expressed (s.v.), and reacting to individual comments (re). The teacher did not introduce any new ideas into the discussion.

The following example taken from the study illustrates the role the teacher took in the class discussions. This dialogue was part of the discussion following 10 minutes of free play in which each pair of pupils had a very immature lemon and a mature lemon.

P1, P2, etc. represent pupil comments.
T represents teacher comments.

P1: The pips in the big lemon were only small.
T: (s.r.) Anyone else want to say something about the pips? (Pause) Yes.
P2: The green one had more than the yellow one. Lots more.
T: (s.r.) Lots more pips in the green lemon? Did other people find that?
P3: We thought that the green one didn't have any pips.
P4: We didn't find any pips in the green one.
T: (re) Didn't you?
P5: We did. We found one and the yellow one had about five.
T: (s.v.) So, you said you found one. Some people said they didn't find any, and someone said they found more. Who was that?
P2: We had ten in the green one and four in the yellow one.
P5: Their green one was a lot more developed than ours. It was a lot bigger.
T: (c.q.) You think that might make a difference?
P2: Yes and as they got bigger the pips die off or something.
T: (s.r.) What do people think about that?
(Laughter)
P6: I think the pips would stay as they get bigger.
T: (s.r.) What do others think?
P7: The pips couldn't die becuase then you couldn't plant the pips to get more lemons.
T: (s.r.) Any other comment?
P8: Yes. Well how come that in some more developed ones there were more pips than in the little ones?
P9: 'Cos in the green ones the little ones were smaller and you could see them developing.
T: (c.q.) You think that the pips were smaller and less developed in the green lemon?
P10: In the green lemon we found that there were all these triangle things and when we dug them out there was a pip in each one.
T: (c.q.) Each triangle had a pip in it?
P11: Or two.
T: (c.q.) Two pips in one triangle?
P11: That's what we found out.
T: (s.v.) We seem to have a difference of opinion. Some people found less pips in the green one than the ripe one. Others found more pips in the green one than in the ripe one. *You* said they could be just developing, and *you* said they might die off as the lemon ripens. Anyone want to say something about this?

(pause)
T: (s.r.) No! Well who ....

Questions for the study

It was decided that the primary questions for the study would be:

(i) Do the provision of materials for a period of free play and the subsequent class discussion both contribute to the number of investigable problems proposed by pupils?
The course planners believe that they do and consideration of the model developed (Symington, 1977) adds support to the course planners' proposition.

(ii) Does the nature of the materials influence the number of investigable problems the pupils propose after the free play or after the discussion? It has been argued that materials which encourage contrasting observations will lead to a greater number of investigable problems being identified.

Since more than one school needed to be used for the study and, since it was desirable to use more than one topic, two secondary questions were considered.

(iii) Is there a difference between schools in the ability of their pupils to propose investigable problems? Crossland's (1967) report leads to the prediction that there will be such a difference where the pupils come from dissimilar home backgrounds or the school climates are not alike.

(iv) Does the number of investigable problems pupils suggest differ from topic to topic? An earlier study (Symington, 1977) found this to be the case.

The design of the study

Two schools situated in a middle class suburb and no more than 2 km apart were chosen and their two grade 6 classes used for the study. In each school the pupils had been assigned to one of the two classes in an almost random manner. The study was carried out in the first two weeks of the school year so as to minimize factors which would lead to differences between the classes. The topics chosen for the study were one from the physical sciences, pendula, and one from the biological sciences, lemons. The researcher acted as the teacher for the programme. The steps taken in each class were identical as follows:

(i) The pupils were given a six page booklet and asked to write on the first page all the questions they could think of about the first topic (either pendula or lemons) which they and a group of friends could investigate using simple everyday materials. There was no time limit.

(ii) The pupils were given 10 minutes "free play" with the materials provided.

(iii) They were asked to record on the second page any additional questions they had thought of that they and a group of friends could investigate using simple everyday materials. Again there was no time limit for the task.

(iv) The teacher conducted a 10 minute class discussion about the phenomena noticed during the period of free play. The discussion was tape recorded. The teacher used only the verbal moves defined earlier.

(v) Following the discussion the pupils were again asked to record, on a new page, any additional 'problems' which occurred to them as a result of the discussion. Again no time limit was imposed.

(vi) Steps (i) and (v) were repeated using the other topic.

The order in which the topics were taken and the materials provided in the various classes for the period of free play are set out in Table 1.

TABLE 1

Topics and materials for the classes.

		First Topic	*Materials provided for 'free play'*	*Second Topic*	*Materials provided for 'free play'*
School X	Class 1	Lemons	Both immature and mature lemons. Knife	Pendula	Single pendulum on frame. Clock
	Class 2	Lemons	Immature lemons only. Knife.	Pendula	Contrasting pendula (see Figure 2) on frame. Clock.
School Y	Class 1	Pendula	Contrasting pendula (see Figure 2) on frame. Clock.	Lemons	Immature lemons only. Knife.
	Class 2	Pendula	Single pendulum on frame. Clock.	Lemons	Both immature and mature lemons. Knife.

Handling of data

The first step was to read all of the problems suggested and to classify them as investigable or non-investigable problems. The definition used for the classification was as follows: 'An investigable problem is one for which a group of grade 5 or 6 pupils could obtain a meaningful answer by investigation using simple everyday materials and with the aid of the teacher.' This procedure has been carried out over a number of experiments and the reliability and validity of this classification method have been described elsewhere (Symington, 1977).

The main emphasis of the study was on the impact of both the period of free play and the class discussion on the number of investigable problems suggested in each class. However, since pre-test data was available it seemed worth looking for differences between schools and between topics using the individual pupil as the unit of analysis. Accordingly, the number of investigable problems on the topic each pupil suggested before the period of free play was identified for each topic. A two way analysis of variance was carried out using these measures.

The investigable problems suggested were then listed out so that those which emerged for the first time after the period of free play or after the discussion could be identified. The final step in the handling of the data was to listen to the tapes of the class discussions and list the phenomena discussed in each class. These could then be checked against the problems proposed by the pupils after the discussion.

Results

The results of the two-way analysis of variance using the pre-test data are shown in Table 2. These results provide answers to the secondary questions.

TABLE 2

Two-way analysis of variance using 'pre-test' measures.

			School	
			X	*Y*
Topic	Lemons	n	55	59
		$\bar{x}$	2.62	3.85
		s.d.	2.31	2.22
	Pendula	n	55	59
		$\bar{x}$	0.62	1.95
		s.d.	0.90	1.72

Source	*SS*	*df*	*MS*	*F*
Phenomena (A)	216.29	1	216.29	60.2*
School (B)	93.29	1	93.29	26.0*
A x B	0.15	1	0.15	0.04
Error	804.44	224	3.59	
Total	1114.17	227		

*$p < 0.001$

Students suggested a greater number of investigable problems on the topic of lemons than for pendula. The F value ($F = 60.2$, $p < 0.001$) indicates that this difference is statistically significant. The result is consistent with that obtained in the earlier study (Symington, 1977). Pupils in school X did not propose as many investigable problems about these topics as did pupils in school Y. This difference was statistically significant. ($F = 26.0$, $p < 0.001$). The method used for the data collection in this study does not provide opportunity to establish whether school climate or home background are causally linked to this result. It can be observed however that, although only 2 km apart, school X has a number of pupils from a housing commission estate whilst school Y has no children from that area. The schools are staffed in a similar manner and so home background is suspected to be the more likely of the two factors to be linked to the observed difference.

The third F value ($F = 0.04$, n.s.) indicates that there is no evidence of differences between the pupils of the two schools being linked to either of the topics considered.

(ii) The next stage in the handling of the data provided answers to the primary questions. The problems proposed after the period of free play and those put forward after the class discussion were divided into two groups. One group consisted of problems put forward for the first time by one pupil but which has been previously proposed by another pupil (Repeated problem). The other group consisted of problems not previously suggested by that class (Additional problem).

TABLE 3

Number of individual problems suggested at each stage.

		Pendulums						Lemons					
		Contrast			Single			Contrast			Single		
		RP	AP	Total	RP	AP	Total	RP	AP	Total	RP	AP	Total
School X	Initially	-	13	13	-	8	8	-	40	40	-	34	34
	After free play	7	11	18	3	9	12	9	29	38	8	31	39
	After discussion	2	1	3	2	2	4	4	4	8	1	10	11
School Y	Initially	-	28	28	-	23	23	-	56	56	-	55	55
	After free play	8	16	24	10	12	22	6	26	32	17	17	34
	After discussion	3	4	7	5	7	12	1	9	10	6	6	12

RP — Repeated problem — a problem previously proposed by one pupil and now proposed by another.
AP — Additional problem — a problem not previously proposed in that class.

The results of this classification are presented in Table 3. The most striking feature was the consistency of the results. In each class both the period of free play and the class discussion following the free play were associated with pupils proposing repeated and additional problems. The greater increase in the number of investigable problems however followed the period of free play in each instance. X^2 tests were carried out and these achieved statistical significance at the 0.05 level. The tables of values are not included here due to the space required. The same pattern of results occurred for both the lemons and pendula.

The hypothesis that the provision of materials which were believed to encourage contrasting observations would enhance the opportunity for pupils to suggest investigable problems was not supported by the data. (Again tables of values were prepared and X^2 tests applied).

The final analysis was the search for matching between the foci of the discussion and the problems suggested by the pupils after the discussion. In each of the classes there were links discernable between discussion foci and the problems listed immediately afterwards. Space does not permit the detailing of these in this paper but one example is included in Table 4.

Conclusion

The data reported in this paper give strong support to the argument that various procedures will facilitate the type of problem based enquiry proposed in science courses such as Science 5/13 and Victorian Primary Science.

The introduction of materials for a period of free play was associated, in this study, with a very significant increase in the number of investigable scientific problems which pupils could propose. The idea that structuring of the materials by the teacher to maximize con-

TABLE 4

Matching foci of class discussion and problems proposed by pupils after the discussion

Discussion Foci

(i) It was noted that the big pendulum swung slower than the smaller one.*
(ii) It was noted that the big pendulum kept swinging longer than the smaller one.*
(iii) One group of pupils reported that if they had the frame on the sloping desk top the large bob didn't hit the base of the frame but when they moved the frame to the floor the large bob hit the base of the frame.*

*In each instance pupils suggested possible causes for the observations.

Problems suggested

(i) Why does the small pendulum go faster than the big one?
(ii) If you dropped the bobs which would hit the floor first?
(iii) Why did the bob hit the frame on the floor when it didn't on the desk?

(An interesting insight was provided by a pupil who after the discussion did not suggest a problem but wrote – I think we should do an experiment at the front and settle everything.)

trasting observations was however not supported by the study.

The class discussion which followed the period of free play was associated with pupil insight into additional investigable problems. In terms of numbers of additional problems the result was not as dramatic as that following the free play. In fact the matching of the discussion foci and the problems proposed suggests that the influence of the discussion may be to focus attention on a few problems in contrast to the divergence following the free play.

In general, those who favour pupil centred enquiry for primary school science should be encouraged by the results of this study. Although the study focussed on the teacher's role, it was the materials which appeared to be the greatest stimulus to pupil thought. Where the teacher played a more obvious role in conducting the class discussion, the techniques used, apparently successfully, were dependent on the use of teaching techniques to promote an exchange of ideas between pupils rather than the teacher's scientific knowledge. Primary teachers generally lack scientific knowledge and it is important to know whether particular science programmes can be effectively handled by such teachers. The results of this study suggest that the skills necessary to at least begin the pupil enquiry do not depend only on the scientific knowledge of the teacher.

References

BARNES, D. *From communication to curriculum.* Harmondsworth: Penguin, 1976.
CROSSLAND, R.W. Report of an individual study of the Nuffield Foundation Primary Science Project. University of Manchester, 1967.
EDUCATION DEPARTMENT OF VICTORIA. *Curriculum guides primary science A, B and C.* Melbourne: Victorian Government Printer, 1969.
SCIENCE 5/13. *With objectives in mind.* London: Macdonald Educational, 1972.

SYMINGTON, D.J. Primary school pupils' ability to see scientific problems in everyday phenomena. *Research in Science Education*, 1977, 7, 41-50.

Research in Science Education, 1978, 8, 175-182

SCIENTIFIC LANGUAGE AND THE HIGH SCHOOL PUPIL : PART 1

P. Lynch, P. Benjamin, T. Chapman, R. Holmes, R. McCammon,
A. Smith, R. Symons

Introduction

The study of classroom dialogue between teachers and pupils in science lessons has revealed a very heavy emphasis on the acquisition of technical language in the classroom (Wilkinson, 1975) In many instances the dialogue can have almost Pinteresque qualities as a consequence of the communication gap between teacher and pupil. The following oft-quoted extract is from Barnes *et al.* (1971):

> T: You get the white . . . what we call casein . . . that's . . . er . . . protein, which is good for you . . . it'll help to build bones . . . and the white is mainly the casein and so it's not actually a solution . . . it's a suspension of very fine particles together with water and various other things which are dissolved in water.
> P1: Sir, at my old school I shook my bottle of milk up and when I looked at it again all the side was covered with . . . er . . ., like particles and . . . er . . . could they be the white particles in the milk . . . ?
> P2: Yes, and gradually they would sediment out, wouldn't they, to the bottom . . . ?
> P3: When the milk goes sour though, it smells like cheese, doesn't it?
> P4: Well it is cheese, isn't it, if you leave it long enough?
> T: Anyway can we get on? . . . We'll leave a few questions for later.

In the above example the pupils attempt to describe phenomena in terms of their own experience and vocabulary, while the teacher, ignoring these attempts, monitors and directs the dialogue towards key words and phrases which represent accurate technical labels.

These observations are disturbing because of the apparent insensitivity of the teaching process and because pupils would seem to be prevented from thinking for themselves, though the communication gap will obviously vary. The observations are interesting in that they pose many unexplored questions related to the need, purpose and scope of technical language in High School Science.

The present work is concerned specifically with examining the acquisition of a technical *word pool* associated with the *nature of matter.* The latter theme is present in any core syllabus in general science and was chosen because it is perhaps the most important centralising conceptual theme at High School level. Philosophically it embraces the Newtonian vision of the universe which so influences scientific naturalism. At a less abstract level it involves subparadigms (forms and states of matter, its particulate nature, the substructure of atoms, the dimensional characteristics involved) the understanding of which are essential for following many of the key technological debates of our time. Thus the theme has considerable bearing on scientific literacy.

The problem

To examine the ability of High School pupils to recognise simple definitions of concept words associated with the theme: the nature of matter.

To investigate the effect of variables such as sex difference, I.Q. (verbal), socioeconomic status of parents, age and grade on the total test score and on individual items.

Method

Subjects and related variables

Six Tasmanian High Schools (12-16 years) were chosen as being representative of the High School population in Tasmania, taking into account catchment area, urban and rural environment and 'probable' socioeconomic status of parents. Every attempt was made to ensure a representative cross section of the ability range at each grade level. A sample of 1635 pupils was obtained from Grade 7 (481), Grade 8 (504), Grade 9 (384), and Grade 10 (241).

The sample consisted of boys (969) and girls (666).

The I.Q. was obtained from a statewide verbal test administered in Grade 5. Pupils' scores were taken from school records.

The socioeconomic ratings were interpreted from reasonably detailed elicited responses, and a standard seven point scale (Hall & Jones, 1950) was used. The sample is observed to be biased towards boys and contains a slightly greater percentage of children from higher socioeconomic groups than is normal for Tasmanian schools.

The word pool

Sixteen concept words were obtained after examining general science textbooks of the last three decades. The common conceptual words associated with the theme: *the nature of matter* have not changed though there have been some changes in their usage. The words were considered to be associated with specific sub-paradigms, described accordingly in Table 1.

TABLE 1

Concept words	*Description of sub-paradigm*
mass, length, area, volume	the *amount, shape and dimensions* of matter
solid, liquid, gas	the *physical form or state* of matter
element, compound, mixture	the *nature of the substances* which make up matter
atom, molecule, ion	the *particulate nature of* matter
electron, proton, neutron	the *substructure of the atom*

Construction of the test

After scrutiny of the High School texts, each concept word was described as a simple definition (Table 2). Multiple choice items, all of the same form, were constructed (example provided). The items were validated by the appropriate matriculation examiners, six very experienced senior teachers and a similar number of new teachers. A pre-test, conducted at a school not involved in the final sample, resulted in the alteration of some distractors and the re-phrasing of two definitions.

TABLE 2

Concept word	*Definition*
mass	The amount of substance in a body.
length	The distance measured along a line.
area	An amount of surface.
volume	A three dimensional space.
solid	Substance X, always keeps the same shape regardless of its container.
liquid	Substance X has a fixed volume but no fixed shape.
gas	A fixed amount of substance X completely fills any container.
element	A substance composed of only one type of atom.
compound	Two or more different elements combine together chemically to form a new substance called a compound.
mixture	A substance that can be seperated into two or more substances by physical means.
electron	The negatively charged particle found in an atom.
proton	The positively charged particle found in an atom.
neutron	The uncharged particle found in an atom.
atom	All substances are made up of small particles called atoms.
molecule	The smallest piece of any substance that can exist on its own.
ion	A charged particle.

The item sequence in the test was randomised – the items were not grouped as sub-paradigms. An example of a multiple choice test item is as follows:

All substances are made up of small particles called:

A elements
B atoms
C chemicals
D compounds

It will be appreciated from Table 2 that the definitions chosen are, for the most part, acceptable generalisations but in some cases they are clearly associated with other major conceptual terms (e.g. liquid: substance X has a fixed volume but no fixed shape). The distractors were chosen so that there is only one correct answer and not a best answer.

Procedure

The test was given to pupils by their science teachers during August 1977. All pupils had thus completed six months of their respective grades. The completion time was standardised to 20 minutes.

Results and findings in respect to the effect of grade

The test score results by grade are given in Fig. 1 (a,b,c,d,e). Each figure shows the

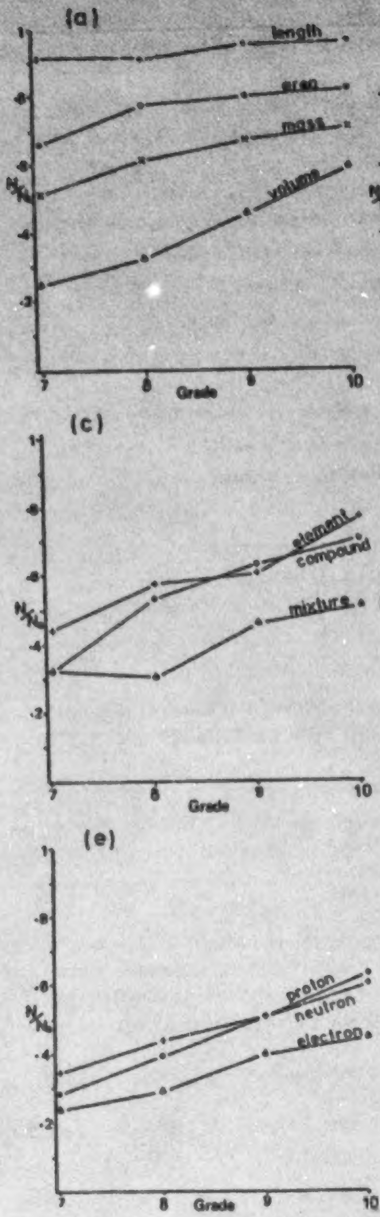


FIGURE 1

Fraction of correct responses, $\frac{N}{N_0}$, by grade for each sub-paradigm test score by grade for definitions associated with the same sub-paradigm.

Science teaching in Tasmanian High Schools is structured and organised by grade. Consequently, it is of particular interest to see how the test score changes with respect to that variable.

For most items there is an increase in test score with grade, though a number of items (length, atom, molecule) show little change over the grade range. Other items appear to reach a plateau after Standard 8 (area, mass solid).

Test scores for length, area, mass and volume are always in the same order and are widely separated at every grade. By comparison gas liquid and solid have very similar test scores for every grade. Test scores for element and compound are very similar at each grade but the score for mixture is significantly lower than the other two after Grade 7. There is quite a dramatic separation of scores for element, compound and mixture after Grade 7.

Atom and molecule show little change in score over the four year High School period with the definition from atom always being more easily recognised than that for molecule. Ion shows a marked and steady increase in score with grade, so that by Grade 10 it achieves the same level as molecule.

Neutron, proton and electron show steady increases in test score with grade. The first two definitions have very similar test scores at every grade but surprisingly that for electron is significantly lower than the other two.

The effect of variables (sex difference, age, grade, I.Q. (verbal), socioeconomic status of parents) on test scores (average total).

The total test score is associated with a mean of 9.0 (S.D. = 3.1). A minimum score of 1 and a maximum of 16 were recorded. Pupils' test scores are well distributed over the total mark range.

(a) All the independent variables chosen produce significant differences (1% level) in average total test scores. Correlation coefficients are provided in Table 3 and are in the order: I.Q. > grade > age > socioeconomic rating > sex difference. The first two variables are clearly the most important. The effect of I.Q. is particulalry pronounced with some individual items, exemplified by the test scores for 'solid', 'liquid', and 'gas'. (Fig. 2).

(b) Sex difference: the average total test score recorded is significantly higher for boys (1% level). Ten of the sixteen individual items are significantly different at the 1% level; nine of these in favour of boys.

(c) Socioeconomic status: The average total test score increases with increase in socioeconomic status.

The effect of these variables on the average total test score was investigated further using stepwise regression analysis. A summary of this analysis is provided in Table 4,

Grade and I.Q. are seen to be the best predictors of score. There are problems of multicolinearity between grade and age, and their respective regression coefficients need to be interpreted with caution (see Table 4). The greatest contribution to the total R^2 (0.28) is associated with I.Q. (0.15) and grade (0.10), and in that order.

TABLE 3

Correlation coefficients between variables

	Total score	*Grade*	*Age*	*I.Q.*	*Socioeconomic status*
Grade	0.3082	-	-	-	-
Age	0.21436	0.80304	-	-	-
I.Q.	0.37769	0.04468	0.18862	-	-
Socioeconomic status	0.13536	0.04770	0.86294	0.12804	-
Sex	0.10986	0.01006	0.02714	0.04893	0.02887

TABLE 4

Summary of regression analysis using all independent variables as predictors of total score

Variable	*B*	*Beta*	*Std Error B*	*F*	*R Square*
Sex	0.85242	0.13470	0.20596	17.130	0.01209
Grade	1.02806	0.32734	0.20786	24.463	0.09862
Age	0.01979	0.00701	0.19003	0.011	0.01279
I.Q.	1.43052	0.38673	0.12660	127.868	0.14662
Socioeconomic status	0.15543	0.10637	0.04830	10.364	0.01090
(Constant)	12.44068				

Overall Multiple R	0.57278
Overall R. Square	0.27802
Adjusted R. Square	0.27382
Standard Error	2.69932
Overall F	52.83417

Discussion

In spite of the fact that all the concept words listed are likely to be treated pictorially, at some stage, involving perhaps a visual aid or model, discussion and application of the concepts still tend to be highly verbal, either in the form of written material or the oral interaction of the classroom. A knowledge and understanding of the concept words would appear to be essential for the understanding of oral and written discuscion associated with most of the important ideas related to the theme in question. Consequently the writer is inclined to consider that the ability to recognise simple definitions of the concept words must also be an important measure of the ability to comprehend the appropriate oral and written communication involving the same words. There are a number of assumptions in this line of argument which we are in the process of investigating. For the moment let us address ourselves to details of the present findings.

The test scores for recognition of individual concept words are not without interest. The dimensional words (mass, length, area, and volume) are associated with a clear, hierarchy of difficulty and 'recognition' is in the same order as Piaget's development of conservation

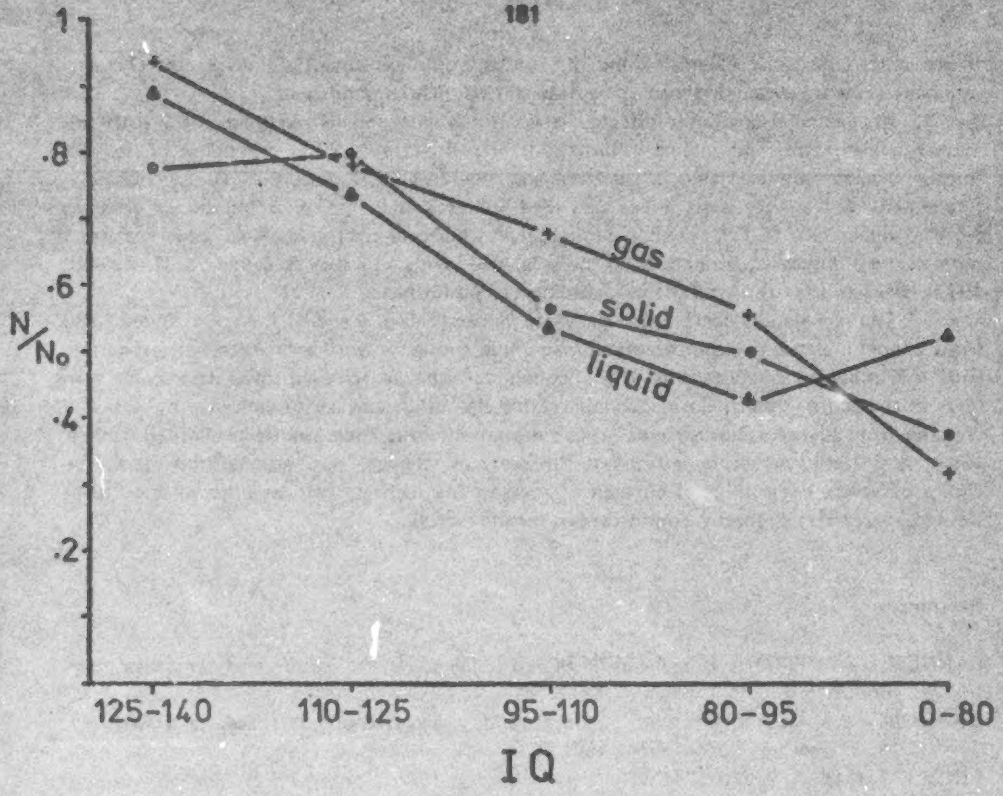


FIGURE 2

Fraction of correct responses, $\frac{N}{N_0}$, by I.Q. for solid, liquid and gas

findings (Ginsberg & Opper, 1969) for these concepts. The scientific description of volume presents many difficulties no doubt partly because of the innate conceptual difficulty associated with the term and partly because of its common usage in other circumstances (e.g., turn up the volume of the radio or T.V., a big book etc.). A study of the free writing of the pupils in this study confirms the latter point.

'Old' concept words, solid, liquid and gas, are all scored significantly lower than the 'newer' word atom for instance. Perhaps the association with the word particle, in the latter case, is something of a give away but it is interesting to note that scores for atom and molecule remain virtually unchanged with grade.

It may come as some surprise to find that the scores for electron are always significantly lower than those for neutron and proton even when all three terms are phrased in a comparable way. However, the pupils' free writing about electrons was mainly concerned with 'electricity' rather than the presence of these particles in atoms, which may offer a clue to an explanation.

Element and compound have similar scores but mixture is scored considerably lower than the other two and does not show a comparable improvement with grade. The latter is a fundamental concept in chemistry and the results would explain why many simple chemical rationalisations are perplexing for the majority of High School pupils.

An examination of the effect of the various independent variables on the total test score provides further interesting information. The total test score is dependent on sex difference, socioeconomic status of parents, age, grade and I.Q., but the last two variables (grade and I.Q.) are the most important. (See Tables 3 and 4). One would expect grade to work in opposition to I.Q., in that pupils tend to perform to a grade level, which acts as a normalising influence thus diminishing the influence of I.Q. (Findley & Bryan, 1971; Daallof, 1971). Even so, the I.Q. is still the best predictor of performance.

The correlation coefficients and Beta values (Tables 3 and 4) for sex difference and socioeconomic status are comparable in magnitude and seem worthy of further investigation. Girls in Tasmanian co-educational High Schools consistently achieved lower test scores than boys in the same schools. Socioeconomic status also influences performance on the test and children from higher socioeconomic groups are advantaged. These results are limited in their scope, and technical vocabulary is not the same as language but, even at this stage, the choice of verbal description of concept words and the teaching emphasis placed upon them would appear to be important considerations for the teacher.

References

BARNES, D., BRITTON, J. & ROSEN, H. (eds). *Language the learner and the school.* Harmondsworth (U.K.): Penguin, 1971.

DAALLOF, U.S. *Ability grouping, context, validity and curriculum process analysis.* New York: Teachers College Press, 1971.

FINDLEY, N.G. & BRYAN, M.W. *Ability grouping 1970: Status, impact and alternatives.* Athens, Georgia: Centre for Education Improvement, 1971.

GINSBERG, H. & OPPER, S. *Piaget's theory of intellectual development.* Englewood Cliffs, New Jersey: Prentice Hall, 1969, 165.

HALL, J. & CARADOG JONES, D. Social grading of occupations. *British Journal of Sociology,* 1950, 1, 38–40.

WILKINSON, A. *Language in education.* London: Oxford University Press, 1975.

Research in Science Education, 1978, 8, 183-193

MISCONCEPTIONS IN PHYSICAL SCIENCE AMONG NON-SCIENCE BACKGROUND STUDENTS

R.D. Linke and M.I. Venz

Introduction

The transition from secondary school studies in non-science areas (or more particularly in areas other than mathematics or the physical sciences) to science-based studies at the tertiary level is a relatively recent phenomenon in Australia, though it is already a substantial and rapidly increasing one. It arises from the relaxation of specific entrance requirements into courses such as Agricultural Science, Biological Science, Earth Science, Engineering, Medicine and the allied Health Sciences; from the broadening of tertiary non-science programmes to allow science-based subjects as elective or enrichment studies; and from the recent development of interdisciplinary courses having no specific subject prerequisites and including substantial components of physical science.

This transition problem involves several closely related aspects. The first is an assumed lack of understanding of basic scientific principles and processes and the consequent development of misconceptions; the second is a possible lack of interest in mathematics and the physical sciences, which might be expected from students who had already rejected these studies at secondary school; and the third relates to the self-perceived inadequacy of these students in their general understanding of science. This leads in turn to three important areas of research: firstly, the diagnosis of both academic deficiencies and related personal difficulties; secondly, the development of appropriate learning materials and teaching strategies; and thirdly, the evaluation of these materials and strategies in relation to the difficulties previously diagnosed.

The present research project was intended to examine the first of these areas. The aim was to investigate the level of understanding among, and nature of misconceptions held by, non-science background students concerning certain fundamental aspects of physical science, and to develop a series of diagnostic tests to identify some of the more serious and more prevalent misconceptions.

Methods

The project was conducted in two stages. In the first stage, a series of interviews was conducted with non-science background students who had enrolled in biological science or medicine at Flinders University or in nursing at Sturt College of Advanced Education. These interviews were structured around questions of basic physical science principles underlying everyday phenomena, for example:

> Where does the heat come from when a candle is burned?
> What happens to salt when it dissolves in water?
> When you comb your hair with a plastic comb, why does the hair tend to cling to the comb?

The purpose of these interviews was to identify the major concept areas in which non-science background students appeared to have serious misconceptions. In the second stage, two of these concept areas were selected as examples with which to establish a method for developing and evaluating a more comprehensive range of diagnostic tests. The areas selected were:

1. Structure of matter, changes of state and solubility;
2. Electricity.

For each area a detailed list of basic principles was defined and the logical inter-relationships between these principles were determined (see Figures 1 and 2). Questions relating to each of the principles were then devised. In most cases duplicate or multiple questions were used to check on the consistency of related misconceptions.

Two tests were constructed for the trial, one concerned with each concept area. Each test was also prepared in two forms, one multiple choice and the other free response. The multiple choice form was intended to provide the basis for the final diagnostic test. The free response form was intended to identify any important misconceptions not covered in the multiple choice form. Each student in the trial was asked to complete one test in multiple choice form and another (from the other concept area) in free response form.

The trial tests were administered to a total of 250 matriculation biology students at four metropolitan high schools and one community college in Adelaide. The schools were chosen to include (but not in any stratified way (students of different ages and from different socioeconomic, ethnic and educational backgrounds. Biology students were chosen because the majority of them had not studied either physics or chemistry at Leaving or Matriculation level (these were classified as non-science background students), yet a reasonable number had concurrently studied both physics and chemistry, providing a useful control group. The total number of students in each science sub-group is shown in Table 1 together with the particular science subjects studied and the concept area in which each group was tested.

TABLE 1

Level of physical science background, subjects studied and tests undertaken by all students

		Physical Science subjects studied at highest indicated level			
		Physics only	*Chemistry only*	*Physics and Chemistry*	
Group I	Grade ≤10	–	-	73	73
MCQ Test: Structure of matter	11	3	7	14	24
	12	6	4	20	30
	Total	9	11	107	127
Group II	Grade ≤10	–	–	63	63
MCQ Test:	11	2	5	12	19
Electricity	12	7	14	17	38
	Total	9	19	92	120

1. Matter exists in 3 different states - solids, liquids, and gases.
   (a) Solids have a definite shape and volume.
   (b) Liquids have a definite volume but no definite shape.
   (c) Gases have no definite volume or shape.

2. All matter is made up of atoms. There are many different types of atoms.

3. All atoms consist:
   (a) A central nucleus containing protons and neutrons.
   (b) A cloud of electrons moving around the nucleus.

   Electrons are much smaller than protons and neutrons.

4. Protons are positively charged; electrons are negatively charged; neutrons have no charge.

5. There is a force of attraction within an atom between the positively charged nucleus and the negatively charged electrons.

6. The number of protons in an atom is the same as the number of electrons; the atom has no overall charge.

7. (a) Elements are substances in which all the atoms have the same number of protons.
   (b) Isotopes are atoms of the same element which have different numbers of neutrons.

8. Atoms of the same or of different elements may group together to form molecules.

9. Compounds are substances in which the molecules are composed of two or more different types of atoms in fixed proportions.

10. There are forces of attraction existing between the atoms of a molecule; these forces are stronger than those between molecules.

11. (a) In a solid the forces of attraction between the constituent atoms or molecules are relatively strong and particles tend to be arranged in an orderly way.
    (b) In a liquid the forces of attraction between the constituent particles are weaker and the particles tend to be disordered.
    (c) In gases the particles are free to move relative to each other and are disordered.

12. Matter can change from one state to another when the degree of movement of the particles changes.

13. (a) The change from gas to liquid is termed condensation.
    (b) The change from liquid to solid is termed freezing.
    (c) The change from solid to liquid is termed melting.
    (d) The change from liquid to gas is termed evaporation.
    (e) The change from solid to gas is termed sublimation.

14. Changes of state for a pure substance, which may be either an element or a compound, occur at fixed temperatures - e.g. water boils at 100°C & freezes at 0°C.

15. The temperature at which changes of state occur depends on the pressure and on the presence of dissolved substances.

16. When a substance changes to a different state in which the particles move more, energy is absorbed. When a substance changes to a state in which the particles move less, energy is released.

17. The heat energy absorbed or given out by a substance during a change of state is termed latent heat. This energy does not affect the temperature of the substance.

18. When a substance changes from one state to another the structure and composition of the atoms or molecules remains the same.

19. (a) A physical change is any change in a substance in which the structure and composition of the particles remains the same.
    (b) A chemical change is a change in which the structure and composition of the particles is altered.

20. When two or more substances are mixed together and the particles of each substance remain the same in structure and composition, the product is called a mixture.

21. When two or more substances are mixed together, the atoms or molecules may react to form new types of particles which are different in structure and/or composition. This process is termed a chemical reaction.

22. There are two special types of mixtures:
    1. (a) A solution is a homogeneous mixture, generally in liquid form, in which the components do not tend to separate out, provided the temperature and pressure remain constant.
       (b) A solution is composed of one or more substances (the solutes) dissolved in another substance (the solvent).

22. 2. A suspension is a heterogeneous mixture, generally in liquid form, in which the components may tend to separate out.

23. Changes of state in a mixture occur at the temperatures of the transition for each constituent substance.

24. Dissolving a solute in a solvent is a physical change and the solute can be recovered from solution, e.g. by heating.

25. (a) Liquids that dissolve in each other are termed miscible.
    (b) Liquids that do not dissolve in each other are termed immiscible.

26. The amount of solute dissolved in a given quantity of solution is the concentration of that solute in solution.

    When the amount of solute dissolved is increased the solution becomes concentrated, and when the amount of solute dissolved is decreased the solution becomes diluted.

27. Concentration may be measured in the following ways:
    (1) mass of solute/volume of solvent;
    (2) mass of solute/volume of solution;
    (3) volume of solute/volume of solvent;
    (4) volume of molecules of solute/volume of solvent.
    (5) percentage weights.

28. At a given temperature and pressure the solubility of a substance is the total amount of substance that will dissolve in a given quantity of solution.

29. (a) In most cases the solubility of a solid in a liquid increases with temperature.
    (b) In most cases the solubility of a gas in a liquid decreases with increasing temperature.
    (c) In most cases the solubility of a gas in a liquid increases with pressure.

30. When no more solute will dissolve in a solvent at a given temperature and pressure the solution is saturated. (This does not apply to miscible liquids.)

FIGURE 1

Basic principles of structure of matter, changes of state and solubility

1. Atoms contain charged particles.

2. There are 2 types of electric charge, termed positive and negative.

3. (a) There is a force of attraction between particles of opposite charge.
   (b) There is a force of repulsion between particles of similar charge.

4. (a) Equal positive and negative charges are neutralised.
   (b) The total electric charge at a point is the sum of the charges.

5. Coulomb's Law:
   The force between two centres of charge varies directly with the product of the two total charges (q1, q2) and inversely with the square of the distance (d) between them.
   $F \propto \frac{q_1 q_2}{d^2}$

6. (a) There is a tendency for charged particles to move to a centre of opposite charge.
   (b) The work required to move a unit of charge between two centres of different charge is is called the potential difference and is measured in volts.

7. (a) The movement of charged particles constitutes an electric current.
   (b) The current is defined as rate of flow of charge and is measured in amperes.

8. (a) There is an inherent resistance to the movement of charged particles through any medium. This resistance depends on the nature of the medium.
   (b) Resistance is measured in ohms.

9. There are two types of electric current:
   (a) Direct current, in which the flow of charge is always in the same direction.
   (b) Alternating current, in which the flow of charge alternates from one direction to the opposite direction.

10. (a) Substances which have a high intrinsic resistance to the flow of an electric current are termed insulators.
    (b) Substances which have a low intrinsic resistance to the flow of an electric current are termed conductors.

11. The flow of a current through a conductor produces a magnetic field around the conductor.

12. The total resistance across a conductor depends on:
    (a) the nature of the conductor;
    (b) the length of the conductor ($R \propto$ length);
    (c) the cross-sectional area of the conductor ($R \propto \text{Area}^{-1}$);
    (d) the temperature.

13. The charge on a conductor tends to be distributed over the surface of the conductor.

14. Ohm's Law:
    The current flowing through a conductor is the ratio of the potential across the conductor to the resistance of the conductor.
    $I = \frac{V}{R}$

15. (a) When 2 conductors of different potential are separated by an insulating medium, the charge on them tends to accumulate at the adjacent ends or surfaces.
    (b) The total charge which can accumulate depends on the area of the adjacent surfaces, the distance (or the thickness of insulating material) between them and the nature of the insulating material.

16. The relative current flowing through interconnected conductors depends on the way they are connected.

17. The ratio of the charge on each of the adjacent surfaces (Q) to the potential difference across the two conductors (V) is termed the capacitance (C) and is measured in farads.
    $C = \frac{Q}{V}$

18. When a current flows firstly through one conductor and then through another, the conductors are said to be connected in series.

19. (a) The flow of an electric current through a conductor transforms energy which is often released as heat.
    (b) The amount of energy transformed depends on the magnitude of the current flowing, which in turn depends on the potential difference and the total resistance across the conductor.

20. The overall resistance of two or more conductors connected in series is the sum of their individual resistances.

21. The energy lost by an electric current is converted into other forms of energy, such as heat, light and sound, and may be used to do mechanical work.

22. (a) When a current flows through two or more conductors simultaneously they are said to be connected in parallel.
    (b) The total current flowing through two or more conductors connected in parallel is the sum of the currents flowing simultaneously through each conductor.

23. (a) The rate of energy released by an electric current is the product of the potential difference and the current.
    (b) This is termed power, and is measured in watts.
    $P = VI$

24. The overall resistance of conductors connected in parallel is given by
    $\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}$
    where R_1 and R_2 are the resistances of the individual conductors.

FIGURE 2

Basic principles of electricity

Results

The following results are discussed in relation to three different sub-groups of students, the first having studied physical science (including general science, physics or chemistry) to grade 10 level or less, the second having studied at least one physical science subject to grade 11 (leaving) level, and the third having studied physics or chemistry at grade 12 (matriculation) level. Where particular numbers or percentages of students are referred to the total (including all three sub-groups) is given as a separate figure followed by the number or percentage for each respective sub-group in the relevant category; thus 22% (14, 17, 47) would indicate a general total of 22% of students, with 14% of those in sub-group 1 (grade 10 science or less), 17% of those in sub-group 2 (grade 11 science) and 47% of those in sub-group 3 (grade 12 science). Students in the first sub-group only are sometimes referred to as non-science background students.

For convenience the results are discussed in relation to four different areas of knowledge or ability in which some of the more pronounced differences between science and non-science students occurred. This is not to suggest that these were the only important areas tested; it simply provides a useful structure for examining the general patterns of student response.

Knowledge and understanding of scientific terms

The terms examined here included such things as 'atom', 'molecule', 'element' and 'isotope' in relation to the structure of matter, as well as 'solubility', 'concentration', and several terms relating to electricity.

Although most students knew that matter was made up of atoms few had any idea of dimensions, and 17% (22, 21, 20) thought that atoms were larger than molecules. Many students also confused atoms with molecules, for example adding the number of symbols in the chemical formula for water to estimate the number of water particles in a single drop. A similar mistake was made by some in estimating the number of particles in a grain of sugar (see Table 2).

TABLE 2

Students' estimates of the number of constituent particles in a grain of sugar and a drop of water

Type and number of Constituent particles	*In a grain of sugar*	*In a drop of water*	*In both a grain of sugar and a drop of water*
Molecules:			
Millions	32% (16, 36, 66)	49% (22, 44, 76)	27% (11, 32, 66)
Thousands	28% (33, 20, 21)	28% (34, 28, 14)	16% (19, 12, 30)
One	13% (15, 12, 10)	9% (7, 12, 10)	5% (4, 8, 3)
Atoms:			
Millions	19% (26, 16, 3)	14% (21, 12, 0)	11% (15, 12, 0)
Thousands	6% (5, 3, 0)	7% (11, 4, 0)	4% (6, 4, 0)
One	2% (1, 4, 0)	2% (4, 0, 0)	-

A number of non-science students knew the relative sizes of atomic particles but were unable to describe the atom. On the other hand, some students described the atom correctly but did not know the relative sizes of the particles referred to. There were some confused descriptions of atoms occurring in the free response test: atoms were described as, for example, 'a ball-like object that circles the proton', 'a centre called a proton with electrons spinning around it fast', and 'round and may be of different size – the oxygen atom is white'. Some students simply recalled diagrams they has seen of the atom 'a round ball with circles around it'.

In describing the characteristics of elements there were 22% (14, 17, 47) of students who realised that all atoms of elements had the same number of protons, and another 29% (16, 50, 43) who thought that the atoms were exactly alike. Isotopes were correctly defined by 46% (19, 71, 90) of students, but 25% (38, 12, 3) defined them as atoms having the same number of component particles but different charges, and 21% (33, 12, 0) as atoms with different chemical properties. There were 21% (33, 12, 0) who thought that carbon dioxide was an isotope of carbon and oxygen, although 6% (7, 8, 0) of these had correctly defined an isotope in the earlier question.

Some students had clearly tried to learn definitions by rote and had simply confused them, for example describing elements as 'atoms with the same number of protons as electrons, that is, neutral charge', and as particles which 'cannot be broken down further: one single atom of a substance'. Similarly, isotopes were defined as atoms with 'the same number of protons but a different number of electrons', and with 'a different number of neutrons – different atomic weight'. Other definitions had no obvious explanation: for example, an isotope was defined as 'the wavelength of a moving particle' and as 'an area that a swinging material is able to swing in equal distances', and an element as an 'object which is sometimes liquid and sometimes solid'.

Terminology relating to solubility was relatively well understood. There were 68% (62, 62, 87) of students who correctly defined solute, solvent and solution, and only 9% (7, 25, 0) could not also define a saturated solution. However the definitions of miscibility and inmiscibility were not so well understood, with 15% (18, 17, 7) of students giving the same definition for both. Overall almost half the students, but only a third of the non-science students – 49% (34, 50, 83) – correctly defined concentration and gave the meanings of concentrated and dilute solutions.

Inconsistency occurred in selecting units for concentration and solubility. A number of non-science students thought that concentration was measured in grams/minute, although only a quarter of these students had previously chosen a definition of concentration related to time. Overall, 76% (67, 83, 90) of students chose the correct unit for concentration but only 60% (47, 58, 93) chose the correct definition.

In relation to electricity, the units of measurement of electrical quantities were generally not well known. Most of the matriculation science students, but only 27% (3, 21, 71) of students overall chose the correct units of measurement for potential difference, current and resistance. Most of the non-science background students thought that current was measured in volts and that potential difference was measured in ohms.

Understanding of scientific concepts and principles

The types of concepts and principles examined here concerned conservation of mass and volume, atomic and molecular forces, changes of state, and characteristics of an electric current.

There were 56% (49, 46, 77) of students who clearly grasped the notion of conservation of volume for both liquids and solids, but only 8% (3, 0, 27) of these also realized that a gas would tend to fill all available space. In all there were 34% (34, 33, 33) of students who thought that volume remained constant for solids, liquids and gases. There were 13% (16, 8, 7) who thought that the volume of a flattened ball of plasticine would be much larger than the original ball, and 44% (4, 8, 0) who thought that the volume would be slightly larger. However only 2% (3, 0, 0) of students also suggested that the volume of a liquid spilled over the floor would be greater than the original volume.

With respect to changes of state, some students thought that the forces within molecules, rather than those between molecules, would decrease when solids were melted, and some thought that the molecules in a solid would show no movement at all. There were 8% (11, 8, 0) who thought that the movement would be constant in all states, and 7% (7, 17, 0) who thought that the attraction between molecules would increase when a solid was melted, but none of these was consistent in suggesting a lower degree of movement.

The general concept of latent heat was not well understood. Only 24% (14, 29, 43) of students knew that the temperature of a melting block of ice would remain steady at 0°C until all the ice had melted. Only 17%. (16, 12, 23) selected the correct definition of latent heat, and 9% (4, 21, 10) confused it with that of specific heat. Although there were 50% (44, 42, 70) of students who realised that the heat absorbed during fusion provided the energy for increased movement of particles, 5% (7, 4, 0) failed to realise in another related question that increased movement would actually occur. In addition there were 26% (41, 25, 13) who thought that heat would change the structure of particles during melting.

Some students were consistent, although incorrect, in their understanding of the properties of a boiling solution. Thus 8% (8, 12, 3) thought that a solution would boil at the average of the boiling points of the constituents, 6% (10, 4, 0) that the whole solution would boil at the lowest constituent boiling point, and 2% (1, 0, 3) at the highest. In the free response test some students explained these latter responses by suggesting that a boiling mixture of alcohol and water 'is diluted but still retains its original properties'. These students clearly had some understanding of the processes involved even though their answers were incorrect.

With respect to the characteristics of an electric current, several students did not realise the difference between a current flowing in a wire and that in a solution (see Table 3). Of the 61% (51, 68, 74) who realised what particles were involved in an electric current in both cases, 32% (19, 68, 50) chose an accurate definition for current as the rate of flow of particles and 14% (0, 21, 42) also knew the correct unit of measurement.

TABLE 3

Students' responses concerning the nature of an electric current in a wire and in solution

Nature of particles in Current flow	*Current in wire*	*Current in salt solution*	*Current in both wire and salt solution*
Electrons	88% (78, 84, 100)	14% (14, 0, 21)	17% (11, 0, 21)
Ions	11% (5, 16, 0)	66% (57, 80, 74)	
Molecules		17% (24, 21, 3)	

The factors influencing the resistance of a conductor were very poorly understood. Only 6% (0, 5, 16) of the students correctly predicted how the resistance of a wire would change if the length or diameter of the wire were changed. Very few non-science students, but most matriculation science students – a total of 33% (21, 16, 55) – correctly predicted the changes in resistance that would occur on varying the length, while 42% (70, 47, 21) thought that resistance would be independent of length. Only 9% (3, 5, 21) of students realised the correct relationship between the resistance of a wire and its diameter, while 9% (3, 11, 18) correctly calculated the proportional change but thought that the resistance of a thicker wire would be greater than that of a thinner wire. Many students thought that resistance was related to diameter rather than to cross-sectional area.

The properties of series and parallel circuits were also poorly understood. Only 13% (5, 21, 13) of students had a working understanding of series circuits, and only 5% (2, 0, 13) had an understanding of parallel circuits. Many students could not work out the change in current that would occur if the total resistance of a wire were changed, although this is a direct application of Ohm's Law.

Ability to apply scientific concepts and principles to common examples

This area concerned the ability of students to recognise and classify common examples of changes of state (e.g. evaporation of sea water, formation of clouds and rain, etc.) and of electrical conductors and insulators (metal, wood, plastics, oil, etc.), and to apply basic definitions concerning, say, electrical charge and forces, to different situations.

Although most of the students knew the terms for the various changes of state (evaporation, condensation, etc.) less than half correctly classified the corresponding examples. Sublimation caused some confusion; many non-science background students seemed to use this term when unsure of the correct response.

A number of students who confused the given examples of physical and chemical changes did so because of lack of understanding of the processes involved. Overall almost half the students – mainly those of non-science background – classified distillation and dissolving as chemical changes. In the free response test these answers were explained in the terms that, for example, during the distillation of alcohol from wine, 'alcohol is produced by a chemical reaction' and that 'molecules change'.

In the questions relating to forces between electrically charged particles, most students knew whether like and unlike charges would repel or attract each other, although many were inconsistent when describing in other questions the type of force acting between similarly and oppositely charged ions. These questions, however, also involved the calculation of overall charge on the ions. In the free response test various suggestions were made concerning the types of forces acting between charged particles. Some students confused forces with bonding and referred to covalent and ionic forces; some referred to positive and negative forces (with the occasional qualification of positive attraction and negative attraction); and a few students chose terms which, although incorrect, conveyed a vaguely appropriate meaning. For instance, one student referred to a 'centrifugal force' instead of the anticipated response of repulsion, and to a 'magnetic force' instead of attraction. Some students considered electrical and magnetic forces to be synonymous, and one referred to a 'repulsive magnetic force'. Other types of forces suggested were 'gravitation', 'cohesion' and 'electromagnetism'.

Many students who stated that certain categories of substance (e.g. solids, liquids, gases) were all conductors or insulators were inconsistent when referring to particular substances within these categories. For instance, of the 7% (2, 5, 3) of students who stated that all solids were conductors and the 6% (6, 11, 3) who stated that all solids were insulators, none consistently classified all the listed solids. Similarly, only 2% (3, 0, 0) of the 37% (48, 32, 21) of students who thought that all liquids were insulators, and none of the 6% (11, 0, 0) who thought that all liquids were conductors, consistently classified all of the listed liquids. Overall about half the students (including only a third of the non-science students) realised that solids and liquids could be either insulators or conductors and correctly classified the more common substances. Many students were not familiar with the conducting properties of gases. Only a few students consistently applied incorrect generalizations, and these were mainly non-science students. In general it seemed that the classification of conductors and insulators was based more on intuition than on clear knowledge of what an electric current was.

Application of basic mathematical skills

This section was concerned with the students' ability to translate certain stated physical laws (including Coulomb's Law, Ohm's Law and the expression of electrical power as a function of voltage and current) into appropriate algebraic functions, to transpose these functions, and to substitute in them simple numerical values. Many students did not seem to possess these basic mathematical skills, as is shown by the summary of results in Table 4.

TABLE 4

Results of questions relating to basic mathematical skills: Translation from a stated law to an algebraic formula; transposition of, and substitution in the algebraic formula

Result	*Coloumb's Law*[1]	*Ohm's Law*[2]	*Power Function*[3]	*Both Ohm's Law and Power Function*
Correct translation	55% (38, 47, 87)	85% (83, 89, 87)	51% (38, 37, 79)	32% (22, 21, 68)
Correct transposition		52% (43, 58, 68)	36% (60, 26, 63)	
Correct substitution		49% (37, 51, 68)	33% (22, 16, 61)	
Interpreted product as				
1. sum	22% (29, 32, 5)	8% (11, 11, 3)	2% (5, 0, 0)	
2. difference			11% (14, 11, 5)	
3. quotient			24% (38, 39, 5)	
Interpreted square as square root	16% (25, 21, 5)			
Interpreted inverse proportion as direct proportion	12% (21, 5, 3)			

1. $F \alpha \frac{q_1 q_2}{d^2}$ 2. $I = \frac{V}{R}$ 3. $P = VI$

Other questions were concerned with simple calculations of ionic charge, concentration, and the interpretation of temperature/time graphs.

A number of students could not manipulate directed numbers. Of the 55% (30, 84, 82) who knew the correct charges on sub-atomic particles, 35% (14, 47, 63) correctly calculated the total charge on two ions, and 7% (13, 18, 5) calculated neither charge correctly. In addition there were 7% (11, 5, 0) who knew that neutrons were uncharged but confused the charges of electrons and protons, but only 2% (5, 0, 0) of these were consistent in calculating ionic charge.

There were 60% (47, 58, 93) of students who correctly defined the term concentration, but only 42% (26, 46, 77) who correctly calculated the concentration of a single solute. A further 8% (5, 12, 10) added the weights of two solutes (coffee and sugar) when calculating the concentration of the second solute (sugar).

The interpretation of graphs was difficult for many students. Although 24% (14, 29, 43) correctly indicated that the temperature of a melting block of ice would remain steady at 0°C, 8% (7, 17, 3) did not choose the corresponding graph, and 25% (34, 21, 37) chose the correct graph without realising that the temperature remained steady. In relation to another question concerning the temperature changes during the melting of butter, 45% (34, 40, 56) of students chose the correct graph in the multiple choice test but very few (approximately 10%) were able in the free response test to sketch the correct graph themselves.

General comments on student responses

There were two purposes in designing these tests. The first and major purpose was to identify certain basic misconceptions in physical science among students whose formal background in this area was very limited. The second was to demonstrate possible trends in the level of understanding of these concepts between students with different levels of physical science background. With respect to the latter purpose there were substantial differences in achievement in both tests between science and non-science background students, with a progressively higher performance according to the level of physical science studied; and in relation to the former purpose there were identified misconceptions, systematically applied across two or more related questions, for almost every basic principle defined. In general, however, these occurred among only a few students, usually — but not exclusively — of non-science background, and the more common characteristic by far of non-science background students seemed simply to be lack of knowledge or confusion rather than systematic misconception. The same characteristic (or it might have been carelessness) was also apparent in many of the responses of matriculation physical science students.

It is tempting to consider these errors and inconsistencies, especially of non-science students, solely as a manifestation of ignorance among students who had, in effect, rejected physical science at the early secondary level, perhaps because of a relatively poor understanding already apparent at that stage. But this is a very superficial interpretation and ignores the fact that many of these students had at least an intuitive understanding of most of the principles tested and were in some cases no more inconsistent than those who had studied science at higher levels. Indeed, of the small numbers of systematic errors revealed in most of the concepts tested, it was the non-science background students who seemed to be more consistent in their application of specific and identifiable misconceptions.

While the number of such students overall was very small it must be remembered that the sample selected was from the matriculation year and included many students who would not be continuing to tertiary studies. It is possible, therefore, that the proportion of these students could be somewhat higher among the smaller group who continue and change their academic emphasis to tertiary science-based studies. On the other hand, the exploratory interviews with tertiary non-science background students conducted during the first stage of the project revealed the same pattern of readily changeable opinions characteristic of the present results. A more reliable estimate of the occurrence of systematic misconceptions among the tertiary group should therefore emerge from a further application of the tests developed here.

In conclusion, it is important in teaching elementary science to adult students to know two things: firstly the nature of general difficulties, whether confusion or misconception, and secondly to identify the individuals, however few, whose particular problem is one of basic and consistent misconception. It has been demonstrated in these trials that both functions can readily be provided within the same test, although there is clearly still a need for further development. This will continue in the third and final stage of the present project.

Research in Science Education, 1978, 8, 195-203

A COMPARISON OF CHANGES IN CHILDREN'S CONCEPTS OF LIFE WITH THE DEVELOPMENT OF RELEVANT CRITERIA IN AUSTRALIAN SCIENCE CURRICULUM MATERIALS

P.P. Sedgwick, R.D. Linke and *A.M. Lucas*

Introduction

Although the concept of life has been investigated by numerous researchers throughout the past 40 years it has usually been within the framework set by Piaget (1929) in which children were classified into one of several stages according to their reasons for attributing life to various objects, and in particular their tendency to attribute life to inanimate objects (animism). For example in a study by Russell and Dennis (1939) using a standardised interview procedure the following stages were identified:

At the lowest level, no concept was evident of what was 'alive'.
Stage I: anything useful or in good condition was considered 'alive'.
Stage II: anything that moved was considered 'alive'.
Stage III: anything showing spontaneous movement was considered 'alive'.
Stage IV: only plants and animals were considered 'alive'.

Other studies have usually given results compatible with these (see Looft & Bartz, 1969 for a review).

Of course movement is not the only attribute to have been investigated. Looft (1974), in a study of second grade children, used the characteristics of respiration (breathing), intake of nutrients and reproduction, and found that none of the children could correctly ascribe all three characteristics to the sixteen objects used. Nutrition was recognised more often than breathing, and breathing more often than reproduction. Smeets (1974) also found that certain 'life traits' such as dying, growing, feeling, hearing, knowing and talking were often attributed to inanimate objects and noted that the reasons given by children often reflected an inadequate understanding of these traits.

However most of these studies have used familiar objects such as cars, chairs, pets (and other common animals), stones and so on, and it is possible that increasing familiarity with the stimulus objects could have inflated the final test scores (Huang & Lee, 1945; Nass, 1956; Klingberg, 1957; Berzonsky, 1971). Moreover the wording of the questions (e.g. 'Does a frog breathe, or need air?'; 'Does a chair need food, or nutrition?' – Looft, 1974) could have led students to consider criteria that they might not otherwise have applied.

In the present study an attempt was made to overcome these difficulties by presenting students with a photograph of an unfamiliar object and asking them to write down as many ways as they could of finding out whether it was alive. This provided more direct evidence than that obtained in most previous studies of the criteria used intuitively by children in determining whether or not an object was alive.

In addition, a survey was made of primary school science curriculum guides used in each State and of some of the more commonly used materials for junior secondary science (particularly in South Australia), to determine whether there was any clear or consistent pattern in the introduction of relevant criteria that might correspond with the changes anticipated (both from this and from previous studies) in the criteria used by children at different grade levels. It was hoped that this might give some indication of the impact of these materials on the development of particular concepts and reasoning skills in children.

Methods

The test was based on a black and white photograph of a two-coloured piece of moistened dough which had been moulded into a ribbed convex shape and placed on beach sand. The photograph was given to each of the children with the following written instructions:

> The object shown above was found on the beach. It was about as large as a person's hand. It was smooth and wet. When picked up it felt soft and spongy.
> How could you find out whether the object was alive? Write down as many ways as you can think of.

In administering the test the following verbal instructions were also given to help the children focus their ideas:

> What would you look for?
> What would you do to it?
> What would you expect it to do?

Ample time was allowed for each child to complete the test – in most cases 10 to 12 minutes was sufficient and the longest time taken was 15 minutes.

The test was administered to a total of 944 students from grades 2 to 10 inclusive at six primary and three secondary schools in the Adelaide metropolitan area. The number of students in each grade and the number of schools represented are shown in Table 1. Intact classes were used (sometimes with mixed grades) and although neither the schools nor classes were selected at random there was no reason to suspect any systematic bias in the selection procedure. All of the tests were conducted during November 1977.

TABLE 1

Composition of the student sample

Grade	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*Total*
Number of schools represented	1	4	4	4	4	3	3	3	3	9
Total number of students	9	125	126	121	105	124	116	110	108	944

In analysing the results a sample of about 100 answer sheets was examined and the responses from these were used as a quide in preparing a detailed classificatory scheme on which to base the final analysis (see Appendix I). The scheme was arranged into six major categories (expert advice; external structure; internal structure and composition; physiological functions; behaviour; and irrelevant answers), with further subdivisions allowing a total of 532 possible categories in which to classify the children's responses.

Every answer sheet was examined and every response classified. Where a single sentence contained more than one conceptually distinct idea it was coded in each of the appropriate categories. For example the statement 'See if it had any mouth, eyes, nose or ears and see if the eyes blink or it can hear and move away from sounds' (13 year old boy, grade 7) was classified as 203, 204, 202, 209, 512, and 571. The reliability of the coding was checked by recording a random sample of 50 answer sheets. There were some difficulties in calculating a coder reliability index since the codes for multiple ideas in a single sentence could be entered legitimately in a different order and there were occasions when there was a different number of codes appearing in the original and recodes. These difficulties are illustrated in the two cases below:

Case 1	Original Code	:	214	251	590		
	Recode	:	214	590	251		
Case 2	Original Code	:	591	623	590		
	Recode	:	590	501	623	550	

In the coder reliability calculations, case one was considered to have three responses in complete agreement. In case two it was assumed that the statement originally coded as 590 was coded as two responses (590 and 501) in the recode, giving one response (623) in complete agreement, one response where the first and last digits were in agreement (590 and 550); one response in agreement in the first and second digits (591 and 590) and one response in complete disagreement (501). The first sentence in case two was 'pick it up and see what it does if it is alive it would move'. In the original coding this was treated as a single response, but in the recode each part was coded separately. There is conceptual agreement between the codings, but they were not scored as agreements in the analysis. Most cases where there were different numbers of responses recognised were similar to this, and the index of agreement is biased toward 'disagreement'. There was complete agreement between first and second codings in 71% of responses; 89% agreement on the major category (first digit); and 80% agreement when the first and third digits only were considered.

Results

Student responses

The results reported here (Table 2) are based upon the substantive reasons for 'living' given by respondents. The second digit codes have been ignored (except in the case of spontaneous movement) for these are generally concerned with a student's method of obtaining information.

Curriculum analysis

A summary of the relevant topics or concepts presented in each of the State Education Department primary school curriculum guides is presented in Table 3. These are classified according to the same basic categories as those derived from the student responses (Appendix I).

In the junior secondary science materials, two of those examined (Australian Science Education Project and Junior Secondary Science Project) did not specifically discuss the question of what a living organism was, but did contain some information on the characteristics of living organisms similar to that presented in the primary school curriculum guides. In the other two resources examined, *In Search of Science* (Russell, Cusack & Mayfield, 1968)

TABLE 2

Percentage of students using category at least once, by grade

Category of response	Code[1]	2	3	4	5	6	7	8	9	10	χ^2 8df
1. Use expert advice	1xy	22.2	8.0	15.1	28.1	16.0	13.7	6.9	16.6	5.6	36.58
No person named	1x0	0.0	1.6	5.6	6.4	1.9	3.2	1.7	5.5	0.9	11.79
Doctor	1x1	0.0	0.0	0.8	1.7	0.0	0.0	0.0	0.9	0.0	7.85
Father	1x2	0.0	2.4	2.4	3.3	0.0	0.8	0.0	0.0	0.0	13.51
Mother	1x3	11.1	2.4	1.6	2.5	0.0	0.8	0.0	0.0	0.0	18.26
Policeman	1x4	11.1	0.0	0.0	0.0	0.0	1.6	0.0	0.0	0.0	42.25
Book	1x5	0.0	0.0	0.0	4.1	1.9	3.2	2.6	4.5	0.9	12.74
Scientist	1x6	0.0	2.4	5.6	8.3	4.8	0.0	2.6	3.6	0.0	20.27
Teacher	1x7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.9	0.0	7.59
Someone who knows	1x8	0.0	0.8	0.8	5.0	4.8	4.8	0.9	2.7	2.8	11.18
Other named person	1x9	0.0	0.8	2.4	4.9	2.9	1.6	0.9	0.9	0.0	11.67
2. Use external structure	2xy	55.6	49.8	46.8	51.2	61.9	63.7	44.0	50.9	50.0	16.10
No specified structure or whole organism	2x0	0.0	20.8	20.6	20.7	26.7	27.4	11.2	16.4	4.6	32.58
Similar to known object or organism	2x1	0.0	13.6	4.0	6.6	7.6	3.2	3.4	5.5	4.6	34.78
Breathing structures (e.g. nose gills)	2x2	0.0	2.4	2.4	5.8	8.6	9.7	15.4	7.3	12.0	26.84
Mouth	2x3	11.1	11.2	11.9	12.4	17.1	19.3	12.9	21.8	23.1	13.23
Eyes	2x4	44.4	16.8	19.8	10.7	17.1	24.2	12.1	17.5	13.0	16.85
Limbs	2x5	22.2	4.8	7.9	2.5	6.7	7.3	6.9	6.4	5.6	8.45
Feelers, tentacles etc.	2x6	0.0	0.0	0.8	4.1	11.4	6.5	5.2	2.7	5.6	25.20
Sex organs	2x7	0.0	0.0	0.0	0.0	0.0	0.8	0.9	2.7	0.0	13.67
Organism inside shell	2x8	0.0	0.8	0.8	2.5	0.0	0.0	0.0	0.0	0.0	12.12
Other named structure	2x9	11.1	8.0	9.5	8.3	13.4	10.5	8.7	11.8	14.0	4.49
3. Use internal structure and/or chemical composition	3xy	0	14.4	15.9	28.1	27.6	31.5	42.2	52.7	57.4	96.84
Structure not specified	3x0	0.0	8.0	9.5	16.5	18.1	16.1	20.7	32.7	29.7	41.74
Cells	3x1	0.0	0.0	1.6	1.7	1.0	0.8	4.3	12.7	21.3	94.01
Subcellular organelles	3x2	0.0	0.0	0.0	0.0	1.0	0.0	1.7	0.9	1.9	8.23
Organic molecules (DNA, portein, etc)	3x3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.9	7.75
Heart	3x4	0.0	2.4	1.6	5.0	2.9	2.4	7.8	1.8	1.9	12.28
Blood	3x5	0.0	4.0	3.2	5.0	6.7	12.9	11.2	12.7	18.5	34.16
Brain	3x6	0.0	0.0	0.8	0.0	0.0	0.0	0.9	1.8	0.9	6.78
Bones/Skeleton	3x7	0.0	0.0	0.0	0.8	1.0	2.4	0.9	1.8	0.9	6.23
Babies	3x8	0.0	0.0	0.0	0.8	1.0	0.0	0.0	0.0	0.0	6.41
Other named organs	3x9	0.0	1.6	0.0	8.3	0.0	3.2	8.6	8.2	12.0	37.35
4. Physiological function	4xy	22.2	32.0	57.1	43.8	56.2	61.3	54.3	53.6	59.3	35.44
Unspecified physical function	4x0	11.1	12.8	11.9	6.6	9.6	8.9	6.9	10.9	12.0	5.21
Heart beat/pulse	4x1	0.0	8.0	22.2	19.0	24.8	36.3	17.2	22.7	16.7	39.80
Breathes	4x2	11.1	16.0	31.7	24.8	26.7	30.6	33.6	15.5	25.9	18.26
Excrete/secrete	4x3	0.0	0.0	1.6	5.0	1.9	2.4	5.2	5.5	8.3	17.37
Temperature regulation	4x4	0.0	0.8	0.0	0.0	1.0	2.4	4.3	4.5	6.5	20.82
Photosynthesis	4x5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.8	23
Respiration/metabolism	4x6	0.0	0.0	0.0	0.8	0.0	0.8	1.7	0.9	8.3	41.22
Reproduction	4x7	0.0	0.0	1.6	0.0	0.0	0.8	5.2	5.5	0.9	24.94
Growth	4x8	0.0	0.0	0.8	1.7	1.9	0.0	6.9	4.9	9.3	35.11
Other specified functions	4x9	0.0	0.0	0.8	1.7	2.9	0.8	3.4	0.9	7.4	26.79
5. Behaviour	5xy	88.9	85.6	92.9	91.7	96.2	97.6	99.1	100.0	95.4	36.83
Unspecified behaviour response	5x0	33.3	38.4	38.1	49.8	61.0	56.5	47.4	56.4	56.5	25.99
Non-specific movement	5x1	44.4	60.0	66.7	70.2	75.2	82.3	89.7	86.4	82.4	51.17
Specific movement	5x2	22.2	38.4	32.5	35.5	50.5	50.8	32.8	32.7	25.0	28.81
Makes a noise	5x3	11.1	4.0	13.5	14.9	11.4	7.3	6.9	10.9	6.5	14.41
Attack defence	5x4	33.3	24.0	12.7	16.5	17.1	14.5	12.1	16.4	5.6	19.78
Changes colour	5x5	0.0	0.8	1.6	2.5	4.8	8.9	6.0	4.5	5.6	15.66
Dehydrates	5x6	0.0	0.0	2.4	2.5	3.9	2.4	2.6	6.4	4.6	10.24
Decays/diseased	5x7	0.0	0.0	2.4	1.7	0.0	1.6	1.7	1.8	0.9	5.12
Eats drinks	5x8	22.2	9.6	10.3	13.2	21.0	18.5	21.6	24.5	19.4	18.16
Mates	5x9	0.0	0.0	0.8	0.0	0.0	0.0	0.0	0.9	0.0	6.05
Spontaneous movement	511 512	0.0	4.8	7.9	15.7	13.3	22.6	19.0	15.5	15.7	24.72

Critical values for χ^2 with 8 degrees of freedom: $p = 0.05$; $\chi^2 = 15.51$
$p = 0.01$; $\chi^2 = 20.09$
$p = 0.001$; $\chi^2 = 26.12$

[1] Categories x and y indicate respectively any of the second and third order alternatives listed in Appendix I.

TABLE 3

Analysis of Curriculum Guides

STATE	CONCEPT CHARACTERISTICS	OUTLINE OF RELEVANT CONCEPTS IN CURRICULUM GUIDES					
		GRADE 2	GRADE 3	GRADE 4	GRADE 5	GRADE 6	GRADE 7
New South Wales	External Structure	Common plants and animals	Birds, Plant growth cycles, Human sensory	Animal sensory organs			(No Primary Grade 7)
	Internal Structure/ Composition		Backbones				
	Physiological Functions	Observation of mice, tadpoles	Need for energy; plant reproduction, eggs.	Reproduction in plants, need for heat, light	Production and storage of food in plants	Regulation of body temperature	
	Behaviour	Ways of moving, eating, drinking, defending	Types of animal movement.	Self protection among animals	Sensitivity in plants; Reproduction in animals and plants		
Victoria	External Structure	Observation of pet animals. Growth and observation of plants	Insects, animals and humans - mouth, legs, eyes, feelers, coverings.	(Syllabus combined with Grade 3)	(Syllabus combined with Grade 3)		(No Primary Grade 7)
	Internal Structure/ Composition	Muscles	Use of magnifying glass and microscope; skeleton, ectoskeleton, blood.				
	Physiological Functions	Requirements for food, water, air. Heartbeat, growth, reproduction	Growth, reproduction; pulse, heartbeat.				
	Behaviour	Movement; daily cycles, response to stimuli	Eating; defence				
Queensland	External Structure	Common plants and animals	Structure of animals determines home, food and habits	Common animals	Types of plants		
	Internal Structure/ Composition					The cell as a basic building block	Chlorophyll in plants
	Physiological Functions			Reproduction in animals	Propagation of plants	Excretion; respiration; reproduction	Photosynthesis; respiration; need for light, water, nutrients in plants
	Behaviour					Movement; responses to stimuli	
South Australia	External Structure	Insects, pet mammals, aquarium creatures	Spiders, insects, tadpoles and frogs	Growing plants; mammals, birds	Types of insects, lizards, plants	Small mammals; sex differences; insects; plants; tadpole and frog	Insects; plants
	Internal Structure/ Composition				Leaf structure in plants; the cell as basic building block		
	Physiological Functions	Life cycles of insects, aquarium creatures	Germination of seeds; need for water	Transpiration in plants	Evaporation, transpiration Insect life cycles	Breathing - gills and lungs; Metamorphosis; maintaining body temperature	Germination of plants
	Behaviour	Observation of insects, aquarium creatures	Tadpole/frog transition	Exercise; care of young. Behaviour of birds	Insect and animal movement - feeding, responses to stimuli	Behaviour of mice and insects, care of young	Decay in plants
Western Australia	External Structure	Common plants and animals	Common plants and animals	Life cycles of plants and animals	(Syllabus combined with Grade 4)	Parts of a plant	(Syllabus combines with Grade 6)
	Internal Structure/ Composition						
	Physiological Functions	Observation of growth and development in plants and animals	Plant and animal development	Growth, development and reproduction in plants		Plant growth and reproduction	
	Behaviour	Eating, drinking, construction of shelters	Eating, drinking; construction of shelters	Responses to stimuli; eating, drinking, sheltering		Animal responses to food, light, heat; learning; migration	
Tasmania	External Structure	Shapes and coverings of plants and animals. Human sensory organs	Parts of the flower; insects, compound eyes	Types of plants; beaks and feet of birds; the human ear			(No Primary Grade 7)
	Internal Structure/ Composition			Skeleton, lungs, brain, heart, kidneys, ear, nerves	Vertebrate and invertebrate animals	Chlorophyll in plants; muscles in animals	
	Physiological Functions	Human senses, need for air, water, food, warmth	Germination	Organ function in animals	Osmosis	Transpiration; digestion; temperature regulation; human senses	
	Behaviour	Observation of birds - some friendly, some shy		Insect movements - hopping, flying, crawling	Migration		

and *Science for Secondary Schools* (Heading *et al.*, 1968), the question 'What is a living organism?' was specifically discussed, and although the emphasis was different from that in the primary school curriculum guides the type of information presented was similar.

Discussion

Student responses

The results presented in Table 2 are broadly consistent with previous studies. The high reliance on behavioural responses, particularly movement, has been commonly reported and the increase in the use of spontaneous movement is consistent with the trend used by Russell and Dennis (1939) as the basis for their scale. This emphasis on behaviour is also consistent with Piaget's early account of children's conceptions of life:

> Although a number of children used such ideas as being able to speak or having blood, etc., to define life, not a single case was found . . . of a child who failed to bring in also the idea of activity and movement. (1960, p. 204)

However the large number of physiological and structural responses has not been consistently reported (this may be an artefact in reporting, with more emphasis being placed on the major criteria). The relative importance of 'blood', and the more frequent use of 'pulse' or 'heart beat' than of 'heart' is not reflected in earlier studies. This may reflect either the difference in methodology between this and other investigations or the debate in the popular press concerning the use of cessation of heart-beat as a criterion for human death.

It is also clear that technical criteria become more common in the higher grades, presumably as a result of specific instruction. The use of 'cells', 'metabolic respiration', 'photosynthesis' and 'temperature regulation' are examples.

Although not specifically coded it is clear from reading the answer sheets that the majority of students have animals, rather than plants, in mind when thinking of life. Although this set could have been encouraged by the verbal instruction 'What would you expect it to do?' it is unlikely that this is the whole explanation.

There are indications in the data that the major change in the pattern of categories used by students occurs at the transition from primary to secondary schooling (grade 7 to grade 8 in South Australia). Whether this is a consequence of the curriculum pattern, of maturity of students or of some other factor cannot be determined from these results.

Curriculum provisions and implications

The amount of science taught in the primary school classes in this sample varied considerably. None of the teachers followed closely the South Australian syllabus, although most used it as their major curriculum resource. In one school the Western Australian syllabus was predominantly used, and in another no syllabus was used at all. Similarly, in the early secondary grades there was a variety of science resources used and none of these was followed systematically. It is not surprising, therefore, that there was little correlation between the introduction in these curriculum guides of particular concepts outlined in Appendix I and their occurrence in the student responses. Nevertheless it is possible to make some general comparisons.

One example concerns the use of movement as a criterion for living. Non-specific movement (e.g. 'see if it moves') was listed by an increasing proportion of students from grade 2 to grade 9, with a slight reduction in grade 10. By contrast the criterion of specific movement (e.g. 'touch it and see if it crawls away'; 'look to see if it has legs and can run') was on the whole mentioned more by primary students than by secondary students. This may reflect the stronger emphasis in the primary curriculum guides on direct observation of a variety of living organisms, looking in particular for different and specific types of movement such as walking, crawling, sliding, flying, etc. as a basis for classifying the organisms. However neither the difference in curriculum emphasis nor the difference in proportion of students using the criterion was very substantial, and this argument must therefore be taken as a fairly tentative one.

In the category of external structure, the response 'no structure specified or whole organism' (e.g. 'examine it to see if it looks like a jellyfish') was mentioned more by all primary grades (except grade 2) than by any of the secondary grades. This again could reflect the stronger emphasis in the primary curriculum guides on observation of whole organisms rather than of particular parts. 'Breathing structures' and 'mouth', for example, were both mentioned more frequently by students in higher grades, though again the difference was not dramatic.

It was in the category of internal structure that the greatest increases in frequency of criterion use over progressive grade levels occurred – from 0% at grade 2 to 57.4% at grade 10. This follows the gradual shift in emphasis in curriculum guides from observation of organisms and their life cycles to the study of internal structures. For example in the South Australian curriculum guide the internal structure of plants is introduced only in grade 5 (see Table 3), and at this level the proportion of students referring to various aspects of internal structure and composition increased quite markedly. A more dramatic example is the reference to cells, which occurred to any noticeable degree only in the secondary grades. This concept did not feature at all in the South Australian primary curriculum guides and was seldom referred to in those of any other State.

Physiological functions were not often mentioned by students in grades 2 and 3 but were mentioned by more than half the students (57.1%) in grade 4. The point of interest here is that in grade 4 the South Australian primary science guide introduces the first detailed study of a mammal including various external structures (mouth, nose, ears, etc.) and their associated functions. These topics are further explored in the grade 6 curriculum guide (see Table 3), where again there was in this study a noticeable increase in the proportion of students referring to physiological functions as criteria for life.

There is no definitive evidence in any of these examples of the impact of curriculum guides on the development in primary or secondary school students of particular concepts of life or on their use of relevant criteria for distinguishing living from non-living. However there is some correspondence between the introduction of certain concepts in curriculum guides and their frequency of use by students that suggests some influence of schooling in this process, especially in relation to the more sophisticated ideas of organic, cellular and molecular function, but to a lesser extent also in the simpler ideas of structure and overt behaviour. How much these ideas are influenced by the school curriculum, and how this influence occurs, are far more difficult for later research to answer.

References

BERZONSKY, M.D. The role of familiarity in children's explanations of physical causality. *Child Development*, 1971, **42**, 705-715.

HEADING, K.E.G. *et. al. Science for secondary schools* (Volumes 1-3, 2nd ed.). Adelaide, South Australia: Rigby, 1968.

HUANG, I. & LEE, W.H. Experimental analysis of child animism. *Journal of Genetic Psychology*, 1945, **66**, 69-74.

KLINGBERG, G. The distinction between living and not living among 7-10 year-old children, with some remarks concerning the so-called animism controversy. *Journal of Genetic Psychology*, 1957, **90**, 227-238.

LOOFT, W.R. Animistic thought in children: Understanding of 'living' across its associated attributes. *Journal of Genetic Psychology*, 1974, **124**, 235-240.

LOOFT, W.R. & BARTZ, W.H. Animism revived. *Psychological Bulletin*, 1969, **71** (1), 1-19.

NASS, M.L. The effects of three variables on children's concept of physical causality. *Journal of abnormal and social psychology*, 1956, **53**, 191-196.

PIAGET, J. *The child's conception of the world.* London: Routledge & Kegan Paul, 1960. First published Harcourt, Brace; New York, N.Y., 1929.

RUSSELL, L.D., CUSACK, J.C. & MAYFIELD, J.M. *In search of science* (Books 1-3). Melbourne: Heinemann Educational, 1968.

RUSSELL, R.W. & DENNIS, W. Studies in animism: I. A standardised procedure for the investigation of animism. *Journal of Genetic Psychology*, 1939, **55**, 389-400.

SMEETS, P.M. The influence of MA and CA on the attribution of life and life traits to animate and inanimate objects. *Journal of Genetic Psychology*, 1974, **124**, 17-27.

Appendix I

CRITERIA USED BY STUDENTS FOR CLASSIFYING LIVING AND NON-LIVING THINGS

MAJOR CATEGORY

1. Expert Advice	Institution	Person
	0 No institution named	0 No person named
	1 Home	1 Doctor
	2 Hospital	2 Father
	3 Laboratory	3 Mother
	4 Library	4 Policeman
	5 Museum	5 Reference book
	6 School	6 Scientist
	7 University/ College	7 Teacher
	8 Place where it would be identified	8 Someone who knows
	9 Other named institution	9 Other named person

3. Internal Structure and Composition	Method of Observation	Type of Structure or Composition
	0 No technique specified	0 Not specified
	1 Look inside	1 Cells
	2 Dissect, cut	2 Subcellular organelles (nucleus, cytoplasm, etc)
	3 Force open, squash	3 Organic molecules (DNA, protein, etc)
	4 Use of microscope	4 Heart
	5 Use of x-ray device	5 Blood, circulating fluid
	6 Chemical analysis	6 Brain
	7 Feed to other organisms	7 Skeleton, bones
	8	8 Babies
	9 Other relevant technique	9 Other named organs (lungs, gut, etc)

5. Behaviour	Type of Stimulus	Type of Response
	0 No technique specified	0 Non-specific response
	1 Explicitly spontaneous	1 Non-specific movement
	2 Generalised non-specific stimulus	2 Specific movement
	3 Change environmental conditions (heat, light, etc.)	3 Makes a noise
	4 Place in different environment (water, sand, etc)	4 Attack/defence
	5 Touch, poke, hit etc.	5 Change in colour
	6 Speak to it	6 Dehydrates
	7 Make a noise	7 Decays or becomes diseased
	8 Feed it	8 Eats, drinks
	9 Other specified stimulus	9 Mates

MAJOR CATEGORY

2. External Structure		Method of Observation/ Comparison	Type of Structure
	method	0 Method or comparison not specified	0 No structure specified or whole organism
		1 Observation	1 Similar to known object or organism
		2 Turn over	2 Breathing structures (nose, gills)
		3 Break open (shell)	3 Mouth
	comparison	4 Idefinite organism	4 Eyes
		5 Animal	5 Limbs
		6 Plant	6 Feelers/tentacles
		7 Inanimate object (clay, stone, etc)	7 Sex organs
		8	8 Organism inside shell
		9 Other relevant object or method	9 Other named structure

4. Physiological Functions	Method of Observation	Type of Function
	0 Technique not specified	0 Unspecified
	1 Listen	1 Pulse, heart beat
	2 Feel, touch	2 Breathing
	3 Look	3 Excretion, secretion
	4 Observation over period of time	4 Temperature regulation
	5 Use instrument or experiment	5 Photosynthesis
	6 Feed it	6 Respiration ($O_2 \rightarrow CO_2$)
	7	7 Reproduction
	8	8 Growth, regeneration
	9 Other specified technique	9 Other functions, including nerve or electrical action

6. Irrelevant Answers		Type of Answer
	0 Not in listed category	0 Avoid - might be harmful
	1 Expert opinion	1 Seek more like it
	2 External structure	2 Reaction of other organisms
	3 Internal structure	3 Irrelevant properties (eg. constant shape)
	4 Physiological function	4 Irrelevant procedures (eg. smell it)
	5 Behaviour	5 Kill it
	6 Combination of above categories	6 Repeating given information
	7	7
	8	8 Obviously facetious
	9	9 Other specified type of irrelevant answer

Research in Science Education, 1978, 8, 205-211

THE PERFORMANCE OF 9 TO 12 YEAR OLD CHILDREN ON PRIMARY, SECONDARY AND MULTIPLICATIVE CLASSIFICATION TASKS: IMPLICATIONS FOR TEACHING AND CURRICULUM DESIGN

Graham W. Dettrick

Introduction

An inspection of primary or junior high school science programmes shows that classification is treated as an important skill through the explicit inclusion of classification experiences or the implicit use of classificatory processes. At the risk of overgeneralizing, classification is 'taught' in early to middle primary grades and 'revised' or 'used' in upper primary or junior high school, but beyond this, a pupil's facility with classification is largely assumed. Much of the 'classification' included in primary and junior high school science programmes refers to the process of forming simple arrangements of objects into subsets that 'belong together'. This single aspect of primary (hierarchical) classification is used almost exclusively. Multiplicative classification (cross classification) surfaces far less frequently and then usually in an application which appears to assume the pupil has control over the process. The process of secondary classification (reclassification) is used rarely either explicitly or implicitly.

Classification occupies a position of primary importance in Piaget's theory of the growth of logical thinking (Inhelder & Piaget, 1969). This prominence is due to several reasons: (1) classification and seriation are the first 'operations' developed by the child (about 6 to 8 years of age), (2) intellectual development during the concrete operational stage is concerned with various aspects of classifying and ordering, (3) an adequate understanding of number must be preceded by an understanding of class inclusion, and, (4) most importantly, the child who can classify can reason logically about the properties of things by adhering to unambiguous criteria. True 'operational' classification means more to Piaget than simply arranging a group of objects into subsets that 'belong together'. An explicit awareness of a number of aspects is necessary which are summarized by intension (properties of similarity and difference), extension (membership by inclusion or exclusion: 'all', 'some', 'one' or 'none'), and the co-ordination of these components. Furthermore, Piaget places considerable emphasis on hindsight and foresight: processes by which a child chooses to revise his earlier actions in the light of those which have followed or where the child anticipates processes which may not be performed until a later time, thereby modifying actions which are carried out in the present.

This study was designed to investigate the ability of upper level primary school pupils in each of three classificatory processes: primary addition of classes, secondary addition of classes and one-to-one multiplication of elements.

Method

Three classification tasks were administered in a randomized order through individual interviews to 108 subjects aged from 9 years 4 months to 12 years 3 months. The subjects were chosen randomly to provide six equal groups according to age and sex. The oral presentation of tasks was tape recorded and transcribed. Silent responses were verbalized by the interviewer to facilitate transcription. The objects used for classification were scale model farm animals manufactured by Britains Ltd. (England).

Tasks

Three Piagetian-type task protocols were designed so that these could be associated with three of the four classificational groupings that Piaget has used as models of logical processes. The tasks were as follows:

Task L_1 (Grouping I): Primary addition of classes – 'Cows and Pigs Task'.
Task L_2 (Grouping II): Secondary addition of classes – 'Collection of Model Objects'.
Task L_4 (Grouping IV): One-to-one multiplication of elements – 'Four x Four Fields Task'.

A brief outline of each task follows. Details of materials, protocol, and scoring criteria used for each task are available from the author.

Task L_1

Each subject was given 12 cows (8 orange, 2 brown and white, and 2 black and white) and 7 pigs (5 pink and 2 black) and was asked (1) to form the animals into groups and give each group a name, (2) whether the spotted ones would be fed if all the cows were to be fed, (3) whether there were more, same, or fewer orange cows than all the cows, (4) whether there were more cows or more animals, (5) whether there would be any animals left if (a) all the cows were taken away and (b) all the pigs were taken away. Supporting reasons for answers to questions 2 to 5 were elicited.

Task L_2

The task used the following model animals:

Standing: One brown horse, one white horse, one brown foal, one white foal, four white sheep (2 eating, 2 not eating), two 'black' sheep (1 eating, 1 not eating), 2 pink piglets (1 eating, 1 not eating).
Lying: One white foal and one brown foal.

Each subject was asked to use the fourteen model animals to form only two groups, name the groups, and give a reason for choosing the name. The interviewer reformed the animals into a single group and the process was repeated with the subject forming several regroupings to investigate whether the secondary classes could be formed and whether a constant superordinate class was maintained.

Additional models not members of the two subsets were introduced to investigate the child's ability to form complementary classes.

Task L_4

The subject was asked to choose from fifteen models provided, which, if any, fitted four empty fields on a four by four multiplicative classification array based on the following properties: cow/calf (big-little), lying/standing, black spotted/brown spotted, and left/right (one way-the other way) facing. Scoring was based on correct choice and elicitation of the properties chosen to support the choice.

Results

For the purpose of statistical analysis, the summarized scoring levels reported in Table 1 were collapsed into pass-fail categories reported in Table 2. Successful completion of all sections of a task was required to 'pass'. The chi-square one-sample test ($\alpha = 0.05$) was used to test for difference of performance on each of the three tasks between (1) age levels and (2) sex of subjects. The results are reported in Tables 2 and 3 respectively. Two of the three calculated chi-square values used to test for differences between age levels exceeded the critical chi-square at the 0.05 level ($L_1 : \chi^2 = 5.77$, df = 1, $L_2 : \chi^2 = 4.87$, df = 1). The hypothesis that there was no difference at the three age levels was rejected for L_1 (primary addition of classes) and L_2 (secondary addition of classes), but the hypothesis was retained for task L_4 (one-to-one multiplication of classes). The chi-square tests for difference of performance according to sex of subjects on each of the three tasks did not exceed the critical chi-square value at the 0.05 level and the hypothesis that there was no difference in performance according to sex was retained.

TABLE 1

Scoring level performance of subjects in each age level for the three classification tasks

		Scoring level						
Age	*Task*	*0*	*1*	*2*	*3*	*4*	*5*	*n*
9;4 – 10;3	L_1	0	1	13	14	8	*	36
10;4 – 11;3	L_1	0	0	4	11	21	*	36
11;4 – 12;3	L_1	0	3	7	11	15	*	36
Total		0	4	24	36	44		108
9;4 – 10;3	L_2	4	5	6	3	1	17	36
10;4 – 11;3	L_2	1	2	9	5	2	17	36
11;4 – 12;3	L_2	2	4	11	7	5	7	36
Total		7	11	26	15	8	41	108
9;4 – 10;3	L_4	9	10	7	5	5	*	36
10;4 – 11;3	L_4	5	8	7	7	9	*	36
11;4 – 12;3	L_4	4	7	8	9	8	*	36
Total		18	25	22	21	22		108

* This level does not apply to this task.

TABLE 2

Results of the chi-square test for significant difference between performance at the three age levels for each of the three classification tasks

	Number of passing scores on each task			
Age	L_1	L_2	L_4	*n*
9;4 – 10;3	8	17	5	36
10;4 – 11;3	21	17	9	36
11;4 – 12;3	15	7	8	36
Total	44	41	22	108
Chi-square	5.77*	4.88*	1.18	
df	1	1	1	

*Significant at the 0.05 level

alpha level = 0.05

TABLE 3

Results of the chi-square test for significant difference between performance on each of the three classification tasks according to sex of subject

	Number of passing scores on each task			
Sex	L_1	L_2	L_4	*n*
Male	20	20	13	54
Female	24	21	9	54
Total	44	41	22	108
Chi-square	0.36	0.24	0.72	
df	1	1	1	

Discussion of results

Significant difference in performance according to age of subjects is restricted to tasks L_1 and L_2. Inspection of Table 2 shows that on task L_1, the youngest subjects (9;4 to 10;2) performed worst and that performance of the oldest (11;4 to 12;3) was worse than the middle group (10;4 to 11;3). Task scores for individual subjects show that few subjects (4%) had difficulty with the criterion of a class (Part A) or were unable to cope with the operation: since cows (B) + pigs (B′) – animals (C) then, B + B′ – C = 0 (null class). A number of subjects (26%) were unable to deal with a problem which required the reasoning: if cows (B) + pigs (B′) = animals (C) then C – B – B′. The greatest difficulty arose with the class inclusion process, that is, if brown cows (A) + spotted cows (A′) = all the cows (B), then A < B or if cows (B) + pigs (B′) = all the animals (C), then B < C. Approximately 59% of the 108 subjects failed to deal consistently with class inclusion relationships. Of the two inclusion relationships in the task, the former (if A + A′ = B, then A < B) proved to be slightly more difficult.

Table 2 indicates that the significant difference between age levels on L_2 was produced by the poor performance of the 11;4 to 12;3 age group in relation to each of the other two groups. A review of L_2 task scores for individuals shows that more subjects in the 11;4 to 12;3 group resisted forming a dichotomy, more subjects did not reclassify, more subjects classified by iterating group membership, and more had difficulty maintaining a constant superordinate class than in the 10;4 to 11;3 age group. Although a review of the performance of the oldest age group is interesting, it does not offer an explanation for the results obtained. Inhelder and Piaget (1969, p. 124) report a similar case of regression in children 10 to 12 years of age for a task which corresponds to Part 3 of task L_2. Inhelder and Piaget's brief comment was 'that regression has nothing to do with the mechanism of classification and simply means that children tend to expect more complicated 'puzzles' than the one which is actually set. (p.124)'. It was the experience of the present investigator that although in a few instances subjects in the 11;4 to 12;3 age group tended to look for more difficulty than there was in problems set, it often appeared that the selection of an inappropriate process form an awareness of a greater range of processes was what led to the regression. It seems that further investigation of classification with subjects near the end of the concrete operational period is necessary to clarify understanding of this regressive behaviour.

An analysis of the performance of the total sample on L_2 shows that 7% of the subjects could not form two groups from the animals presented because they were overwhelmed by the differences. A much larger percentage (17%) failed to place one animal correctly into the L_4 four by four multiplicative matrix for the corresponding reason. These subjects tended to choose animals which were 'different' from the others and support the selection with a reason like, 'Those are all different and you don't have one of these on the board so I think this one (a horse) can go in that empty field and this (an orange/Jersey calf) can go in that empty field'.

Of the 38% of subjects who were classified as 'passing' task L_2 (that is, reading scoring level 5) very few (5%) formed distributions with complementary class names such as eating versus not eating, feathered versus without feathers or lying versus not lying. Most 'successful' responses were made in terms of descriptive opposites, for example, dark versus light colour, lying versus standing, wild versus tame, little versus big, and so on. It was for this reason that responses to Parts 2 and 3 of the protocol were collapsed to form scoring criterion 5. Furthermore, very few subjects were able to operate effectively with, or were even prepared to admit the possibility of, a rearrangement in the association of the models so that the membership of one group or class was limited to a single member (L_2: Part 3).

The finding that there was no significant difference in performance between males and females on each of the three tasks is congruent with the conclusions which follow from the assumption that logico-mathematical groupings are models of the cognitive processes with which the child establishes relationships leading to fully operational logical concepts in the concrete operational period.

If differences were found between task performances on the basis of sex, this would suggest that the tasks or the method of administration measured processes which were not entirely dependent upon logical ability or which were so closely related to certain experiences which are generally socially exclusive on the basis of sex that the experience and social transmission factors, which Piaget considers to be important contributors to intellectual development, were responsible for the difference.

Conclusions: Implications for teaching and curriculum design

The results from the tasks indicate that, by the age of twelve years, children's classificational abilities are far from fully developed. Since the classificational processes examined in this investigation are firmly located in the concrete operational period (Piaget, 1973, p. 23; p. 110-111), it seems clear that this period extends beyond the age usually associated with the end of primary schooling, and may not merge with the development of formal operational thought until the age of fourteen or more years.

The range of classification ability observed within the 9;4 to 12;3 age group may serve as a partial reason why it is justifiable to anticipate that children will exhibit a wide range of aptitude to deal with scientific and mathematical concepts in both primary and junior high school and is strongly suggestive of the need to deliberately structure teaching strategies, expectations, and material experiences to take these differences into account. That is, teaching processes centered around common experiences under conditions of equal exposure to instruction will inevitably produce individual student achievement distributed along a curve that approaches the normal, with achievement scores remaining highly correlated with initial aptitude (Carroll, 1963; Raven, 1970; Bloom, 1971; Dettrick, 1977).

Although it may appear to be expedient to teach concepts through the application of rules for those children who have not established the necessary conceptual processes (Raven, 1970; Youngiss, 1971) the long-term effects of training, in particular that associated with current teaching processes, appears to be diminutive in terms of retention of ability and generalizibility according to the performance of children in this and other studies of classification (Raven, 1957; Kofsky, 1966; Parker & Day, 1971). A more beneficial and lasting effect may be anticipated by recognizing the concrete operational nature of children's thought and the need for the child's continued personal use of a variety of relevant manipulable materials. As teaching strategies and curriculum programmes in mathematics and science move further and further away from concrete physical experiences to the more formal and symbolic, the appropriateness of the teaching strategies and curriculum programmes decrease accordingly for most students up to the age of fourteen years. Furthermore, amelioration of the development of logical processes such as those inherent in classification is more likely to occur if teachers provide problem solving opportunities and use questioning strategies which cause children to go beyond the mere manipulation and arrangement of things that 'go together' over a far greater period of children's schooling than is usual at present. Not only extended involvement with classificational processes, but also practice in a wide range of situations, and most importantly the repeated necessity to reflect upon their logical decision-making may facilitate the improvement of children's classification skills. To advance these developments, the heightening of awareness of students in teacher education programmes, attention to the classificational performance capacities of children and development of strategies to deal with this during in-service education, as well as the establishment of clear and unambiguous guidelines in curriculum programmes appear to be necessary.

Copies of materials used in the teacher education programme at Gippsland Institute of Advanced Education to directly and indirectly heighten the awareness of students are available from the author of this paper.

References

BLOOM, B.S. *Individual differences in school achievement: A vanishing point?* (Phi Delta Kappa Monograph Series). Bloomington, Ind: Phi Delta Kappa Inc., 1971.

CARROL, J. A model of school learning. *Teachers College Record*, 1963, 64, 723-733.

DETTRICK, G.W. Models of schooling and models of evaluation. Proceedings of the Australian Association for Research in Education Conference: Curriculum Evaluation. Canberra, 1977, 1-11.

INHELDER, B. & PIAGET, J. *The early growth of logic in the child.* New York: Norton & Co., 1969.

KOFSKY, E.A. A scalogram study of classificatory development. *Child Development*, 1966, 37, 191-204.

OVERTON, W.F. & JORDAN, R. Stimulus preference and multiplicative classification in children. *Developmental Psychology*, 1971, 5, 505-510.

PARKER, R.K. & DAY, M.C. The use of perceptual, functional and abstract attributes in multiple classification. *Developmental Psychology*, 1971, 5, 312-319.

PIAGET, J. *The child and reality.* New York: Crossman Publishers, 1973.

RAVEN, R.J. The development of classification abilities in culturally disadvantaged children. *Journal of Research in Science Teaching*, 1967, 5, 224-229.

RAVEN, R.J. The effects of a structured learning sequence on second and third grade children's classification achievement. *Journal of Research in Science Teaching*, 1970, 7, 153-160.

YOUNGISS, J. Classificatory schemes in relation to class inclusion before and after training. *Human Development*, 1971, 14, 171-183.

THE CONTRIBUTORS

JOHN AINLEY	Australian Council for Educational Research, Melbourne.
WARREN BEASLEY	Faculty of Education, University of Queensland.
JIM BUTLER	Faculty of Education, University of Queensland.
JOHN DEKKERS	Faculty of Education, James Cook University of North Queensland.
GRAHAM DETTRICK	Gippsland Institute of Advanced Education, Victoria.
ROD FAWNS	Faculty of Education, University of Melbourne.
ADRIAN FORDHAM	Australian Council for Educational Research, Melbourne.
BARRY FRASER	Faculty of Education, Macquarie University.
ELIZABETH HEGARTY	School of Microbiology, The University of New South Wales.
RUSSELL LINKE	Educational Research Unit, The Flinders University of South Australia.
LAWRENCE LITTLE	Sydney Teachers' College.
ARTHUR LUCAS	School of Education, The Flinders University of South Australia.
PADDY LYNCH	Faculty of Education, University of Tasmania.
ROB de M-MACLAY	Sydney Teachers' College.
MAX MADDOCK	Faculty of Education, University of Newcastle.
DAVID MITCHELL	Kelvin Grove College of Advanced Education, Brisbane.
MILES NELSON	Churchlands College of Advanced Education, Western Australia.
BARRY NEWMAN	School of Education, The University of New South Wales.
JOHN OWEN	Melbourne State College, Victoria.
JOHN RENTOUL	School of Education, Macquarie University, New South Wales.

JAMES RICHMOND	Mount Gravatt College of Advanced Education, Queensland.
G. WAYAN SEREGEG	School of Education, La Trobe University, Victoria.
P. SEDGWICK	Educational Research Unit, The Flinders University of South Australia.
DAVID SYMINGTON	Toorak State College, Victoria.

THE CONFERENCE ORGANISERS

To the many persons who assisted in organising and conducting this conference, my thanks. In particular the advice and assistance of the following were invaluable:

Cec Burr	Mount Gravatt College of Advanced Education
Jim Butler	University of Queensland
Ken Kretschner	Coorparoo High School
Herbie Martin	Mount Gravatt College of Advanced Education
David Mitchell	Kelvin Grove College of Advanced Education
Jim Richmond	Mount Gravatt College of Advanced Education

together with Mrs H. Martyn for her secretarial assistance and the technical staff of Mount Gravatt College of Advanced Education.

Cam McRobbie